Research Methods in

PSYCHOLOGY

Tenth Edition

Onuncu Basımdan Çeviri

PSİKOLOJİDE

Araştırma Yöntemleri

John J. Shaughnessy

Eugene B. Zechmeister

Jeanne S. Zechmeister

Çeviri Editörü: Doç. Dr. İlyas Göz

Mc Graw Hill Education

nobel

Psikolojide Araştırma Yöntemleri

NOBEL AKADEMİK YAYINCILIK EĞİTİM DANIŞMANLIK TİC. LTD

Yayın No.: 1550
Psikoloji PDR No.: 62
ISBN: 978-605-320-456-5
© 10. Basımdan Çeviri, Eylül 2016

PSİKOLOJİDE ARAŞTIRMA YÖNTEMLERİ
RESEARCH METHODS IN PSYCHOLOGY
John J. Shaughnessy - Eugene B. Zechmeister - Jeanne S. Zechmeister
Çeviri Editörü: Doç. Dr. İlyas Göz

Copyright 2016, NOBEL AKADEMİK YAYINCILIK EĞİTİM DANIŞMANLIK TİC. LTD. ŞTİ. SERTİFİKA NO.: 20779
Bu baskının bütün hakları Nobel Akademik Yayıncılık Eğitim Danışmanlık Tic. Ltd. Şti.ne aittir. Yayınevinin yazılı izni olmaksızın, kitabın tümünün veya bir kısmının elektronik, mekanik ya da fotokopi yoluyla basımı, yayımı, çoğaltımı ve dağıtımı yapılamaz.

Original edition copyright © 2015 by McGraw-Hill Education. All rights reserved.
Turkish edition copyright 2016 by Nobel Akademik Yayıncılık. All rights reserved.

Genel Yayın Yönetmeni: Nevzat Argun -nargun@nobelyayin.com-
Dizi Editörü: Naim Dilek -naim@nobelyayin.com-

Redaksiyon: Naim Dilek -naim@nobelyayin.com-
Sayfa Tasarım: Emel Yıldız -emel@nobelyayin.com-
Kapak Tasarım: Mehtap Yürümez -mehtap@nobelyayin.com-

Baskı ve Cilt: Atalay Matbaacılık Sertifika No.: 15689
Büyük Sanayi 1. Cad. Elif Sk. no.: 7\236-237 İskitler/ANKARA

KÜTÜPHANE BİLGİ KARTI

Shaughnessy J., John. Zechmeister B., Eugene. Zechmeister S., Jeanne.
Research Methods in Psychology / *John J. Shaughnessy - Eugene B. Zechmeister - Jeanne S. Zechmeister*
Psikolojide Araştırma Yöntemleri / *Çeviri Editörü: Doç. Dr. İlyas Göz*
10. Basımdan Çeviri, XX + 492 s., 170x240 mm
Kaynakça ve dizin var.
ISBN 978-605-320-456-5
1. Psikoloji 2. Araştırma Yöntemleri

Genel Dağıtım: ATLAS AKADEMİK BASIM YAYIN DAĞITIM TİC. LTD. ŞTİ.
Sipariş: -siparis@nobelyayin.com-
Telefon: +90 312 278 50 77 - Faks: 0 312 278 21 65
e-satış: www.nobelkitap.com - esatis@nobelkitap.com
Bilgi: www.atlaskitap.com - info@atlaskitap.com

Dağıtım: Alfa, Arasta, Final, Gökkuşağı Dağıtım, Kida, N-T mağazaları ve Prefix

NOBEL AKADEMİK YAYINCILIK EĞİTİM DANIŞMANLIK TİC. LT
Ankara Büro: Mithatpaşa Cad. No.: 74/4 Kızılay / A
Tel: 0312 418 20 10 **Faks:** 0312 418 30 20
Rasimpaşa Mah. Soğütlüçayır Sok. No.: 16/21 Kat: 5
Kadıköy / İSTANBUL **Tel / Faks:** +90 (216) 449 20 1
nobel@nobelyayin.com - www.nobelyayin.com

www.nobelkitap.con

Çeviri Kurulu

KISIM I: *Genel Meseleler*

1 Giriş
2 Bilimsel Yöntem
3 Psikoloji Araştırmalarında Etik Meseleler

Doç. Dr. İlyas Göz - *Acıbadem Üniversitesi*

KISIM II: *Betimsel Yöntemler*

4 Gözlem
5 Anket Araştırmaları

Yrd. Doç. Dr. Sema Karakelle - *İstanbul Üniversitesi*

KISIM III: *Deneysel Yöntemler*

6 Bağımsız Grup Desenleri
7 Tekrarlı Ölçümler Desenleri
8 Karmaşık Desenler

Doç. Dr. Handan Can - *Uludağ Üniversitesi*

KISIM IV: *Uygulamalı Araştırmalar*

9 Tek Denekli Araştırma Desenleri

Prof. Dr. Sonia Amado - *Ege Üniversitesi*

10 Yarı Deneysel Desenler ve Program Değerlendirme

Doç. Dr. İlyas Göz - *Acıbadem Üniversitesi*

KISIM V: *Araştırmanın Analizi ve Rapor Edilmesi*

11 Verilerin Analizi ve Analizin Yorumlanması: Kısım 1. Verilerin Betimlenmesi, Güven Aralıkları, Korelasyon
12 Verilerin Analizi ve Analizin Yorumlanması: Kısım 2. İstatistiksel Anlamlılığın Test Edilmesi ve Analizin Öyküsü
13 Psikolojide İletişim

Yrd. Doç. Dr. Osman İyilikçi - *İzmir*

İçindekiler

Ön Söz xiii

KISIM I
Genel Meseleler 1

1 Giriş 2

PSİKOLOJİ BİLİMİ 3
BİLİMİN BAĞLAMI 6
Tarihsel Bağlam 6
Sosyal ve Kültürel Bağlam 9
Ahlakî Bağlam 13
BİR ARAŞTIRMACI GİBİ DÜŞÜNMEK 14
Medyada Yayımlanan Araştırma Bulgularını Değerlendirme 17
Araştırmaya Başlama 18
ÖZET 23
TEMEL KAVRAMLAR 24
GÖZDEN GEÇİRME SORULARI 24
DÜŞÜNDÜRÜCÜ SORULAR 25
İLERİ ARAŞTIRMALARIN CEVAPLARI 25
DÜŞÜNDÜRÜCÜ SORULARIN CEVAPLARI 26

2 Bilimsel Yöntem 27

BİLGİYE BİLİMSEL YAKLAŞIM VE GÜNLÜK YAKLAŞIM 28
Genel Yaklaşım ve Tutum 29
Gözlem 30
Kavramlar 33
Rapor Etme 35
Araçlar 36
Ölçme 37
Hipotezler 38
BİLİMSEL YÖNTEMİN AMAÇLARI 40
Betimleme 40
Kestirim 44
Açıklama 46
Uygulama 49
BİLİMSEL TEORİLERİN KURULMASI VE TEST EDİLMESİ 50
ÖZET 54
TEMEL KAVRAMLAR 55
GÖZDEN GEÇİRME SORULARI 55
DÜŞÜNDÜRÜCÜ SORULAR 56
İLERİ ALIŞTIRMALARIN CEVAPLARI 57
DÜŞÜNDÜRÜCÜ SORULARIN CEVAPLARI 57

3 Psikoloji Araştırmalarında Etik Meseleler 58

GİRİŞ 59
ARAŞTIRMAYA BAŞLAMADAN ÖNCE GÖZ ÖNÜNE ALINMASI GEREKEN ETİK MESELELER 60
RİSK/FAYDA ORANI 62
Risk Belirleme 63
En Az (Minimal) Risk 65
Riskin Üstesinden Gelme 65
BİLGİLENDİRİLMİŞ ONAM 67
PSİKOLOJİ ARAŞTIRMALARINDA ALDATMA 73
ARAŞTIRMA SONRASI BİLGİLENDİRME (DEBRIEFING) 76
HAYVANLARLA ARAŞTIRMA 78
PSİKOLOJİ ARAŞTIRMALARININ RAPOR EDİLMESİ 80
ETİK AÇIDAN UYGUN BİR ARAŞTIRMA YAPMANIN AŞAMALARI 83
ÖZET 85
TEMEL KAVRAMLAR 86
GÖZDEN GEÇİRME SORULARI 86
DÜŞÜNDÜRÜCÜ SORULAR 87

KISIM II
Betimsel Yöntemler 91

4 Gözlem 92

GENEL BAKIŞ 93
DAVRANIŞTAN ÖRNEK SEÇME 94
Zaman Örnekleme 94
Durum Örnekleme 96
GÖZLEME DAYALI YÖNTEMLER 97
DOĞRUDAN GÖZLEME DAYALI YÖNTEMLER 97

Müdahale Yapılmayan Gözlem 99
Müdahale Edilerek Yapılan Gözlem 100
DOLAYLI (DİKKAT ÇEKMEYEN) GÖZLEME DAYALI YÖNTEMLER 106
Fiziksel İzler 107
Arşiv Kayıtları 109
DAVRANIŞIN KAYDEDİLMESİ 112
Kapsamlı Davranış Kayıtları 113
Seçilmiş Davranış Kayıtları 114
GÖZLEME DAYALI VERİLERİN ANALİZİ 118
Niteliksel Veri Analizi 118
Niceliksel Veri Analizi 120
GÖZLEME DAYALI ARAŞTIRMALAR ÜZERİNE ELEŞTİREL DÜŞÜNME 124
Gözlemcinin Etkisi 125
Gözlemci Yanlılığı 129
ÖZET 131
TEMEL KAVRAMLAR 132
GÖZDEN GEÇİRME SORULARI 132
DÜŞÜNDÜRÜCÜ SORULAR 133
İLERİ ALIŞTIRMALARIN CEVAPLARI 134
DÜŞÜNDÜRÜCÜ SORULARIN CEVAPLARI 1 134

5 *Anket Araştırmaları 135*

GENEL BAKIŞ 136
ANKETLERİN KULLANIMI 136
ANKET ARAŞTIRMALARININ ÖZELLİKLERİ 138
ANKET ARAŞTIRMALARINDA ÖRNEKLEM SEÇME 138
Örnek Seçmenin Temel Kavramları 138
Örneklem Seçme Yaklaşımları 140
ANKET YÖNTEMLERİ 144
Posta Anketleri 145
Yüz Yüze Görüşmeler 146
Telefon Görüşmeleri 147
İnternet Anketleri 148
ANKET ARAŞTIRMA DESENLERİ 151
Kesitsel Desen 151
Ardışık Bağımsız Örneklemler Deseni 152
Boylamsal Desen 154
ANKET FORMU 158
Bir Araç Olarak Anket Formları 158
Öz Bildirim Ölçümlerinin Güvenirlik ve Geçerlikleri 160
Bir Anket Formunun Hazırlanması 162
ANKET ARAŞTIRMALARI ÜZERİNE ELEŞTİREL DÜŞÜNME 167
Bildirilen Davranış ile Gerçek Davranış Arasındaki Uyum 168
İlişkisellik (Korelasyon) ve Nedensellik 169
ÖZET 171
TEMEL KAVRAMLAR 173
GÖZDEN GEÇİRME SORULARI 173
DÜŞÜNDÜRÜCÜ SORULAR 174
İLERİ ALIŞTIRMA I'İN CEVAPLARI 175
İLERİ ALIŞTIRMA II'NİN CEVAPLARI 176
DÜŞÜNDÜRÜCÜ SORULARIN CEVAPLARI 176

KISIM III
Deneysel Yöntemler 177

6 *Bağımsız Grup Desenleri 178*

GENEL BAKIŞ 179
PSİKOLOGLAR NEDEN DENEY YAPAR 179
DENEYSEL ARAŞTIRMANIN MANTIĞI 180
SEÇKİSİZ (YANSIZ) GRUPLAR DESENİ 181
Bir Seçkisiz Gruplar Deseni Örneği 182
Blok Seçkisizleştirme 188
İç Geçerliğe Tehditler 189
DENEYSEL BULGULARIN ANALİZİ VE YORUMLANMASI 195
Deneylerde Veri Analizinin Rolü 195
Sonuçların Betimlenmesi 197
Verilerin Anlamının Onaylanması 200
Veri Analizi Bize Ne Söyleyemez 204
DENEYSEL BULGULARIN DIŞ GEÇERLİĞİNİ SAĞLAMA 205
EŞLEŞTİRİLMİŞ GRUPLAR DESENİ 209
DOĞAL GRUPLAR DESENİ 211
ÖZET 213
TEMEL KAVRAMLAR 215

GÖZDEN GEÇİRME SORULARI 215
DÜŞÜNDÜRÜCÜ SORULAR 216
İLERİ ALIŞTIRMA I'İN CEVAPLARI 218
İLERİ ALIŞTIRMA II'İN CEVAPLARI 218
DÜŞÜNDÜRÜCÜ SORULARIN CEVAPLARI 218

7 Tekrarlı Ölçümler Desenleri 219

GENEL BAKIŞ 220
ARAŞTIRMACILAR NEDEN TEKRARLI ÖLÇÜMLER DESENLERİNİ KULLANIRLAR 220
TEKRARLI ÖLÇÜMLER DESENLERİNDE ALIŞTIRMA ETKİLERİNİN ROLÜ 222
Alıştırma Etkilerini Belirleme 222
Tamamlanmış Desende Alıştırma Etkilerini Dengeleme 224
Tamamlanmamış Desende Alıştırma Etkilerini Dengeleme 228
TEKRARLI ÖLÇÜMLER DESENLERİNDE VERİ ANALİZİ 233
Sonuçların Betimlenmesi 233
Verilerin Anlamını Onaylama 235
AYRIMSAL AKTARMA SORUNU 237
ÖZET 238
TEMEL KAVRAMLAR 240
GÖZDEN GEÇİRME SORULARI 240
DÜŞÜNDÜRÜCÜ SORULAR 240
İLERİ ALIŞTIRMALARIN CEVAPLARI 241
DÜŞÜNDÜRÜCÜ SORULARIN CEVAPLARI 1 242

8 Karmaşık Desenler 243

GENEL BAKIŞ 244
KARMAŞIK BİR DESENDEKİ ETKİLERİN BETİMLENMESİ 244
Bir 2 × 2 Desen Örneği 245
Temel Etki ve Etkileşim 247
Etkileşimin Açıklanması 251
Üç Bağımsız Değişkenli Karmaşık Desenler 253
KARMAŞIK DESENLERİN ANALİZİ 255
Etkileşimin Olduğu Bir Durumda Analiz Planı 257
Etkileşimin Olmadığı Bir Durumda Analiz Planı 260
ETKİLEŞİM ETKİLERİNİ YORUMLAMAK 262
Etkileşim Etkileri ve Teori Test Etme 262
Etkileşim ve Dış Geçerlik 263
Etkileşim Etkileri ve Tavan ve Taban Etkileri 265
Etkileşim ve Doğal Gruplar Deseni 266
ÖZET 269
TEMEL KAVRAMLAR 270
GÖZDEN GEÇİRME SORULARI 270
DÜŞÜNDÜRÜCÜ SORULAR 271
İLERİ ALIŞTIRMA I'İN CEVAPLARI 272
İLERİ ALIŞTIRMA II'İN CEVAPLARI 273
DÜŞÜNDÜRÜCÜ SORULARIN CEVAPLARI 273

KISIM IV
Uygulamalı Araştırma 275

9 Tek Denekli Araştırma Desenleri 276

GENEL BAKIŞ 277
VAKA ÇALIŞMASI YÖNTEMİ 278
Özellikleri 278
Vaka Çalışması Yönteminin Avantajları 282
Vaka Çalışması Yönteminin Dezavantajları 285
Vaka Çalışmasına Dayanan Tanıklıklar Üzerine Eleştirel Düşünme 287
TEK DENEKLİ (KÜÇÜK-N) DENEYSEL DESENLER 288
Tek Denekli Deneylerin Özellikleri 289
Özel Deneysel Desenler 291
Tüm Tek Denekli Tüm Desenlerin Ortak Sorunları ve Sınırlılıkları 297
ÖZET 300
TEMEL KAVRAMLAR 301
GÖZDEN GEÇİRME SORULARI 301
DÜŞÜNDÜRÜCÜ SORULAR 301
İLERİ ALIŞTIRMALARIN CEVAPLARI 302
DÜŞÜNDÜRÜCÜ SORULARIN CEVAPLARI 1 303

10 Yarı Deneysel Desenler ve Program Değerlendirme 304

GENEL BAKIŞ 305
GERÇEK DENEYLER 306
Gerçek Deneylerin Özellikleri 306
Doğal Ortamda Gerçek Deney Yapmanın Zorlukları 307
Gerçek Deneylerde İç Geçerliğe Tehditler 309
Gerçek Deneylerde Bile Kontrol Edilemeyen Sorunlar 314
YARI DENEYSEL DESENLER 316
Denk Olmayan Kontrol Grup Deseni 318
Denk Olmayan Kontrol Grup Deseni: Langer ve Rodin'in Çalışması 319
Denk Olmayan Kontrol Grup Deseninde İç Geçerliğe Tehditler 321
Dış Geçerlik Meselesi 325
Kesintili Zaman Dizileri Deseni 326
Denk Olmayan Kontrol Gruplu Zaman Dizileri Deseni 330
PROGRAM DEĞERLENDİRME 331
ÖZET 336
TEMEL KAVRAMLAR 337
GÖZDEN GEÇİRME SORULARI 337
DÜŞÜNDÜRÜCÜ SORULAR 338
İLERİ ALIŞTIRMALARIN CEVAPLARI 339
DÜŞÜNDÜRÜCÜ SORULARIN CEVAPLARI 340

KISIM V
Araştırmanın Analizi ve Rapor Edilmesi 341

11 Verilerin Analizi ve Analizin Yorumlanması: Kısım 1. Verilerin Betimlenmesi, Güven Aralıkları, Korelasyon 342

GENEL BAKIŞ 343
ANALİZİN ÖYKÜSÜ 344
BİLGİSAYAR TEMELLİ VERİ ANALİZİ 345
ÖRNEK: ORTALAMALARIN KARŞILAŞTIRILDIĞI BİR DENEY İÇİN VERİ ANALİZİ 346
Aşama 1: Verileri Tanımaya Başlama 347
Aşama 2: Verileri Özetleme 351
Aşama 3: Verilerin Anlamını Onaylanmak İçin Güven Aralıklarının Kullanılması 356
ÖRNEK: KORELASYONEL ÇALIŞMA İÇİN VERİ ANALİZİ 367
Aşama 1: Verileri Tanımaya Başlama 368
Aşama 2: Verileri Özetleme 368
Aşama 3: Korelasyon İçin Güven Aralığının Hesaplanması 373
ÖZET 373
TEMEL KAVRAMLAR 375
GÖZDEN GEÇİRME SORULARI 375
DÜŞÜNDÜRÜCÜ SORULAR 376
İLERİ ALIŞTIRMALARIN CEVAPLARI 377
DÜŞÜNDÜRÜCÜ SORULARIN CEVAPLARI 377

12 Verilerin Analizi ve Analizin Yorumlaması: Kısım 2. İstatistiksel Anlamlılığın Test Edilmesi ve Analizin Öyküsü 379

GENEL BAKIŞ 380
SIFIR HİPOTEZİ ANLAMLILIK TESTİ (SHAT) 380
DENEYSEL DUYARLILIK VE İSTATİSTİKSEL GÜÇ 384
SHAT: İKİ ORTALAMANIN KARŞILAŞTIRILMASI 386
Bağımsız Gruplar 387
Tekrarlı Ölçümler Desenleri 387
İSTATİSTİKSEL ANLAMLILIK VE BİLİMSEL YA DA PRATİK ANLAMLILIK 388
İKİ ORTALAMANIN KARŞILAŞTIRILMASI İLE İLGİLİ ÖNERİLER 389
İKİ ORTALAMAYI KARŞILAŞTIRIRKEN SONUÇLARIN RAPOR EDİLMESİ 390
İKİDEN FAZLA KOŞUL İÇİN VERİ ANALİZİ 392
TEK FAKTÖRLÜ BAĞIMSIZ GRUPLAR DESENİ İÇİN ANOVA 392
Üç veya Daha Fazla Bağımsız Grup İçeren Desenler İçin Etki Büyüklüğünün Hesaplanması 398

Bağımsız Gruplar Deseni İçin Gücün Değerlendirilmesi 399
Çoklu Grup Deneylerinde Ortalamaların Karşılaştırılması 400
TEKRARLI ÖLÇÜMLER VARYANS ANALİZİ 403
BAĞIMSIZ GRUPLAR DESENLERİ İÇİN İKİ FAKTÖRLÜ VARYANS ANALİZİ 407
Etkileşimin Olduğu Karmaşık Desen Analizi 407
Etkileşimin Olmadığı Durumda Analiz 409
Bağımsız Grupları İçeren İki Faktörlü Desene Ait Etki Büyüklükleri 410
KARMAŞIK DESENLERİN ANALİZİNDE GÜVEN ARALIKLARININ ROLÜ 410
KARMA DESEN İÇİN İKİ FAKTÖRLÜ VARYANS ANALİZİ 411
KARMAŞIK DESEN İÇİN SONUÇLARIN RAPOR EDİLMESİ 413
ÖZET 413
TEMEL KAVRAMLAR 414
GÖZDEN GEÇİRME SORULARI 415
DÜŞÜNDÜRÜCÜ SORULAR 415
İLERİ ALIŞTIRMALARIN CEVAPLARI 416
DÜŞÜNDÜRÜCÜ SORULARIN CEVAPLARI 416

13 Psikolojide İletişim 417
GİRİŞ 418
İNTERNET VE ARAŞTIRMA 420
ETKİLİ YAZIM İÇİN ÖNERİLER 421
ARAŞTIRMA RAPORUNUN YAPISI 423
Başlık Sayfası 424
Özet 424
Giriş 425
Yöntem 427
Sonuçlar 427
Tartışma 431
Kaynaklar 432
Dipnotlar 433
Ekler 434
SÖZEL SUNUMLAR 434
ARAŞTIRMA ÖNERİLERİ 435

Ekler: İstatistik Tabloları 437
Sözlük 442
Kaynakça 451
Grafik, Tablo ve Kutucukların Alındığı Kaynaklar 470
İsim Dizini 473
Konu Dizini 479

Yazarlar Hakkında

JOHN J. SHAUGHNESSY Michigan-Holland'daki nispeten küçük, seçkin bir üniversite olan Hope College'da profesördür. Chicago'daki Loyola Üniversitesinden 1969 yılında lisans derecesini aldıktan sonra, 1972 yılında Northwestern üniversitesinden doktora diplomasını aldı. Psikoloji Bilimleri Birliği'nin (Association for Psychological Science) ve Midwestern Psikoloji Birliği'nin (Midwestern Psychological Association) üyesidir. Yeni bir çalışmasında belleğin uygulamalı yönlerine odaklanmıştır. Benton J. Underwood ile birlikte *Experimentation in Psychology* (Wiley, 1975) kitabının yazarıdır. Öğrenciler kendisini 1992 yılında Hope Üniversitesinin en iyi profesörü olarak seçtiler ve yine aynı üniversitede 2008 yılında Janet L. Andersen Excellence in Teaching ödülünü aldı.

EUGENE B. ZECHMEISTER Chicago'da büyük bir üniversite olan Loyola üniversitesinde psikoloji profesörüdür. Bu üniversitede 1970'ten bu yana hem lisans hem de lisans üstü düzeyde dersler vermektedir. New Mexico Üniversitesinden 1966 yılında mezun olmuştur. Yüksek lisans ve doktora diplomalarını ise, sırasıyla, 1968 ve 1970 yıllarında Northwestern Üniversitesinden almıştır. Bilişsel psikoloji ve deneysel yöntem alanlarında uzman olan Profesör Zechmeister bellek, eleştirel düşünme, istatistik ve araştırma yöntemleri üzerine yazılmış olan kitaplarda ortak yazar olarak görev almıştır. Kendisi Amerikan Psikologlar Birliği'nin (APA) ve Psikoloji Bilimleri Birliği'nin (Association for Psychological Science) üyesidir. 1994 yılında Loyola University Sujack Award for Teaching Excellence in the College of Arts and Sciences ödülünü aldı.

JEANNES S. ZECHMEISTER Chicago'daki Loyola Üniversitesinin psikoloji fakültesinde 1990-2001 yılları arasında öğretim üyeliği yaptı. Lisans diplomasını 1983 yılında Wisconsin-Madison üniversitesinden, yüksek lisans diplomasını 1998 yılında ve doktora diplomasını 1990 yılında Northwestern Üniversitesinden klinik psikoloji alanında aldı. Araştırma yöntemleri alanında lisans ve lisans üstü dersler verdi ve bağışlayıcılık (forgiveness) ile bağlantılı psikolojik süreçler üzerinde çalıştı. Bir öğretici olarak gücünü, öğreticilik sıralamalarında uzun yıllar üst sıralarda yer alarak ve her yıl Loyola'da en iyi öğretim üyelerinden birisi seçilerek kanıtladı. Dr. Zechmeister hâlen Charlottesville'de çalışmaktadır.

Paula'ya
(J.J.S.)

Ruth O'Keane,
James O'Keane,
Kathleen O'Keane Zechmeister
ve
Annemin anısına
(E.B.Z.)

Babam Harold W. Sumi'nin
anısına
(J.S.Z.)

Ön Söz

Elinizdeki bu onuncu baskıda, önceki baskılarımızda belirlediğimiz amaçlara ulaşmak için çalışmaya devam ediyoruz: Öğrencilere hem araştırma süreçleri hakkında heyecan verecek hem de onların araştırma yöntemlerinde yetkin birer uygulamacı olmalarını sağlayacak bir giriş yapmak. Konuların iyi bir organizasyonu ve açık yazılmış metinler, yetkinleşmeye yardımcı olabilir, ama araştırma süreçleri konusunda öğrencilerin hevesini kamçılamak ayrı bir konudur. Coşkulu bir öğretim üyesi kilit bir öneme sahiptir ama öğrenciyi heveslendirmeye bizim de katkımızın olabileceğine inanıyoruz. Bizim yaklaşımımız, betimlediğimiz araştırma yöntemlerinin davranış hakkındaki bilginin ilerletilmesi için nasıl kullanıldığını göstererek, öğrenciyi konuların içine çekmektir. Bu amaçla, yeni ve (umuyoruz) ilginç psikoloji araştırmalarından yararlandık; bunu yaparken çeşitli kaynaklardan alıntılar yaptık: Örneğin, *Developmental Psychology; Journal of Personality and Social Psychology; Ethology; Psychological Science; Law and Human Behavior; Science; Journal of Cross-Cultural Psychology; Preventive Medicine; The Sports Psychologist; Cyberpsychology; Behavior & Social Networking; Perspectives on Psychological Science*, bunlardan bazılarıdır.

Bu kitabı yeni kullananlar için temel organizasyonu, pedagojik yardımları, ve başlıca özellikleri tekrar ele alarak başlıyoruz. Önceki baskıları kullanmış olanlar, yeni olan hususları görmek için doğrudan "Bu Baskıdaki Değişiklikler" başlığına gidebilirler.

ORGANİZASYON VE YAKLAŞIM

Yaklaşımımız, uzun yıllar süren öğretim üyeliği deneyimimize dayanıyor. Araştırma yöntemleri öğretim üyeleri olarak biliyoruz ki sınıflarımızdaki öğrencilerin çoğu, araştırma yapan insanlar değil, araştırmaları okuyan insanlardır. Her ikisini de yapan öğrenciler, eleştirel düşünme becerilerini geliştireceklerdir. İnanıyoruz ki öğrencilerimize yapabileceğimiz en iyi yardım, araştırma yöntemleri konusuna problem çözme yaklaşımı benimseyerek, eleştirel düşünmeye yöneltmektir. Newsweek yazarı Sharon Begley'in yakınlardaki bir bilim eğitimi yazısında belirttiği gibi, "Bilim, olguların toplanması değil, dünyanın sorgulanmasıdır." Dahası, "Öğretebileceğimiz en yararlı beceri, kendimize ve başkalarına "*Nerden biliyorsunuz?*" sorusunu sorma alışkanlığıdır" (*Newsweek*, 8 Kasım, 2010; s. 26).

Araştırmacılar iyi bir soru ile araştırmalarına başlarlar ve daha sonra, bu soruyu en iyi şekilde cevaplamalarına yardım edecek bir araştırma yöntemi seçerler. Bazen zahmetli olan veri toplama görevi, araştırma sürecinin sadece başlangıcıdır. Verileri analiz etme ve yorumlama, psikolojik süreçler hakkında iddialarda bulunabilmek için aynı derecede önemlidir. Araştırmacılar (ve öğrenciler), elde ettikleri bulguların niteliğini eleştirel şekilde değerlendirebilmek için, seçmiş oldukları yöntemin güçlü ve zayıf yönlerini analiz etmelidirler.

Son baskıdan devam ettiğimiz bir başka özellik, kitabımızın web sitesindeki tasarımıdır. Öğrencilerin kitaptan öğrendikleri bilgilerini test etmeleri için, diğer

önemli psikoloji web sitelerine bağlantıların yanı sıra, etkileşimli alıştırmalar ve sorular yer almaktadır. Öğretim üyeleri, öğretmen kullanma kılavuzunu ve anlatım/tartışma desteğini yararlı bulacaklardır. Hem öğrenciler hem de öğretim üyeleri, şu siteden yazarlarla kolayca iletişim kurabilirler: *www.mhhe.com/shaughnessy9e*.

Her baskıdaki yaklaşımımızla uyumlu olarak öğrenciler, soruları cevaplamak için çok yöntemli bir yaklaşımı benimsemenin, psikoloji biliminin gelişmesi için en iyi yaklaşım olduğunu öğreneceklerdir ve bu kitabın bir amacı, onları bir araştırmayı yapabilmeleri için gerekli olan stratejilerle donatmaktır. Bu nedenle giriş bölümlerinden sonraki organizasyonumuz, yöntem açısından, en basit gözlem tekniklerinden karmaşık deneysel desenlere ilerlemektedir. Psikoloji araştırmalarında etik meselelere ve araştırmacıların, hayvan veya insan davranışını araştırdıkları zaman, karşılaştıkları ikilemlere duyarlıyız. Bu konudaki duyarlılığımızın bir göstergesi olarak, "Etik" konusuna ayrı bir bölüm ayırdık (Bölüm 3) ama belirli yöntemlerle bağlantılı olduğu zaman, diğer bölümlerde de özel etik sorunlarını tartıştık.

İnanıyoruz ki araştırma yöntemleri en iyi, yayımlanmış olan psikoloji araştırmaları üzerinde öğretilir. Bu nedenle, bizim tartışmakta olduğumuz yöntemleri gerçekten kullanmış olan araştırma örneklerini geniş ölçüde kullandık. Araştırma örneklerini keyifle güncelledik ama öte yandan, araştırma yöntemlerini öğrenmeye yardımcı olduğu gösterilmiş olan araştırmaları ve önemli "klasik" bulguları muhafaza ettik. İnanıyoruz ki bilgi peşinde koşmanın bu heyecanlı yolunda öğrencileri bize katılmaları için motive etmenin bir yolu, psikoloji araştırmalarının sağladığı "doyum"u göstermektir.

Öğrencilerin, kendileri için en önemli olduğunu ve öğrenmeleri gerektiğini düşündüğümüz hususları açık şekilde görebilmeleri için, bölümler içinde madde işaretleri ve bölümlerin sonunda ise Gözden Geçirme Soruları kullanmaya devam ediyoruz. Ve öğrencilerin öğrendikleri prensipleri uygulamalarını öğrenmelerine yardım etmek için, Düşündürücü Sorular kullanmaya devam ediyoruz. Öğrencilerin, bir yandan araştırma prensiplerini öğrenirken, bir yandan da bu prensipleri uygulamalarına olanak sağlamak için Düşündürücü Sorular modelini geliştirerek, çoğu bölüme İleri Alıştırmalar ekledik. Geniş bir istatistik özeti, yine kitabın sonundadır (Bölüm 11 ve 12), ve bu meselelere kitabın uygun yerlerinde kısaca temas etmeye devam ediyoruz. Bunun bir şekli, bizim "İstatistik İpucu" dediğimiz, öğrencilerin dikkatini istatistiksel analiz sorularına çeken bir pedagojik destektir. Bazı durumlarda, bu soruları biz öğrenciler için kendimiz cevaplıyoruz; diğer durumlarda Bölüm 11 ve 12'deki materyale yöneltiyoruz. İnanıyoruz ki bu yöntem, öğretim üyelerinin istatistiği ne zaman ve ne şekilde konuya dâhil edeceklerine kendilerinin karar vermelerine olanak sağlayacaktır.

BU BASKIDAKİ DEĞİŞİKLİKLER

- Her gözden geçirmede olduğu gibi, sunumumuzu daha açık hâle getirmek için çalıştık. Öğrencilerin kavramları daha iyi anlamalarını kolaylaştırmak için cümlelerde ve paragrafların yapısında küçük değişiklikler yapıldı.
- Bilimsel psikolojinin temel yöntemleri yıldan yıla çok az değişir; ama araştırma eğilimleri, veri toplama teknikleri, araştırma bulguları ve önemli

tartışma konuları sürekli olarak değişir. Hem bilimsel psikolojideki önemli çağdaş meselelerin açıklamasını hem de yeni eğilimleri ve teknikleri yansıtan düzinelerle yeni yayını elinizdeki yeni baskıya dâhil ettik. Elinizdeki onuncu baskıya, çoğunluğu 2010 yılından sonra yayımlanmış olan 100'den fazla yayın eklendi.

- Bazıları bir bilim olarak psikolojiyi, araştırmalardaki katılımcıları ağırlıklı olarak Batılı, eğitimli, sanayileşmiş, zengin ve demokratik ülkelerden seçtiği için eleştirdi (Jones, 2010; bk. Bölüm 1). Hatta bu ülkelerde bile kadınlar, azınlıklar ve göçmenler gibi birçok grup yıllarca araştırmalara dâhil edilmedi. Ama her şey değişiyor. *Journal of Latina/o Psychology* gibi süreli yayınlar, psikoloji araştırmalarını bu gruplara götürüyor ve kültürlerarası araştırma eğilimindeki artışı gösteriyor. Örneğin, biz Robert Levine'in kültürlerarası "yardım davranışı" araştırmalarını ve daha sonra kültürler içindeki yardım davranışı üzerine yapılmış olan korelasyonel bir çalışmayı inceleyeceğiz (bk., Bölüm 2'de Knafo, Schwartz & Levine, 2009). Bölüm 4'te, Nairan, Ramirez-Esparza'nın Meksikalı ve ABD'li öğrenciler arasındaki dil farkları üzerine yaptığı araştırmayı betimleyeceğiz. Psikoloji giderek daha fazla uluslararası bir nitelik kazanıyor; bu nedenle, daha çok araştırmadan bahsedeceğiz.
- Günümüzdeki araştırmacılar, internet aracılığıyla milyonlarca katılımcıya ulaşabiliyorlar. Bu olanak, özellikle anket kullanan araştırmacılar için önemli hâle geldi. Bölüm 5'te, internette anket yapmak isteyen öğrencilere önerilerimiz olacak.
- Aynı zamanda, internetteki davranışın kendisi araştırma konusu olmuştur. Sosyal ağ siteleri ve sohbet odaları, çoğunlukla katılımcılar gözlenerek, sosyal psikologlar ve kişilik psikologları tarafından araştırıldı (bk., Bölüm 4). Siber psikoloji alanındaki bilim dergilerinin ortaya çıkması, psikoloji araştırmalarındaki bu yeni eğilimin bir göstergesidir.
- Veri toplama, çoğunlukla incelikli tekniklerle ve araçlarla mümkün olmaktadır. Bunun bir örneği, James Pennebaker, Matthias Mehl, Nairan Ramirez-Esparza ve arkadaşlarının dil davranışı, mutluluk, ve davranışsal sağlık konulu araştırmalarında kullandıkları elektronik kayıt cihazıdır (bk., Bölüm 4). Bölüm 7'de, ilk bakışta katılımcıların romantik arkadaşlarına benzemesi için dijital olarak değiştirilmiş fotoğrafların nasıl cinsiyet farkları ortaya çıkardığını göstereceğiz (Günaydin, Zayas, Selcuk & Hazan, 2012).
- Daha önceki baskılarda olduğu gibi, araştırma girişimlerinde etik endişeleri öne çıkardık; bu amaçla, bir bölümün tamamını bu konuya ayırdık (Bölüm 3); ama belirli yöntemleri işlerken etik konusunu da ele almaya devam ediyoruz. Bu baskıda, araştırmacıların insan katılımcıları korumak için nasıl etik kurallara uygun çalıştıklarını göstermek için, Stanley Milgram'ın ünlü araştırmasını "tekrarlayan" Jerry Burger'in çalışmasına dikkat çekeceğiz. Birçok internet araştırması, ciddi etik sorunlara yol açmaktadır. EAR gibi elektronik olarak harekete geçen teknikler kullanarak yapılan gözlemler de böyledir ve bunları da tartışacağız.

Söz etmeye değer diğer küçük değişiklikler ise şunlardır:

- Bölüm 1'de, sadece *etnik merkezcilik* demek yerine, *etnik merkezci yanlılık* ifadesini kullandık.

- Bölüm 6'da eşleştirilmiş gruplar desenlerini tartışırken, *eşleştirme görevi* yerine *eşleştirme değişkeni* terimini kullandık.
- Bölüm 9'da, çağdaş kullanımla (örneğin, Kazdin, 2011) paralellik sağlamak için *tek denekli desen* yerine *tek denekli araştırma deseni* terimini kullandık. Bu baskıda Bölüm 9'un adı *Tek Denekli Araştırma Deseni'dir* ve, önceki baskılarda olduğu gibi, hem vaka çalışmalarını hem de tek denekli deneyleri ele almaktadır. Bu bölümün, öğretim üyelerinin ders programına her zaman giremeyeceğinin farkındayız. Ama bu bölümdeki konuların önemli olduğuna inanıyoruz çünkü özellikle birçok üniversite öğrencisi uygulamalı psikoloji alanında çalışmak istiyor. Nedensel çıkarım yapılırken bir tek vakaya dayanmanın birçok tehlikesine işaret ettik; bu durumla, özellikle araştırma bulguları popüler medyada yayımlandığı zaman karşılaşılmaktadır.
- Bu baskıda, öğrencilerin araştırma konularına karşı meraklarını uyandırmaya yarayacak bilgiler içeren yeni "kutular" dâhil ettik. Örneğin, Bölüm 9'da Stacy Lopresti-Goodman ve arkadaşlarının araştırmasını ele alacağız. Bu araştırma, evde şempanze beslemenin ve yaban hayvanı eti ticaretinin neden olduğu yetim şempanzelerin uğradıkları psikolojik yıkımı belgelemektedir. Ve Bölüm 10'da, istatistiksel regresyonu "Haber Olma Uğursuzluğu" bağlamında açıklayacağız.
- Langer ve Rodin'in (1976) bir bakımevi ortamındaki sorumluluk değişimlemesini içeren klasik yarı deneysel çalışması, Bölüm 10'da yarı deneysel desenler tartışmamızın temeli olmaya devam ediyor. Ama buna, 11 Eylül terör saldırısı sonrasını konu alan yeni zaman dizileri desenleri (Peterson & Seligman, 2003) ve bir kentte sigarayı yasaklamanın sağlığa etkilerini konu alan araştırmaları (Khuder ve ark., 2007) ekledik. Bu bölümdeki program değerlendirme tartışmamız, sağlık sigortası gibi büyük ölçekli sosyal programları ele alıyor.
- Nihayet, araştırma örneklerinin güncellenmesi ve her bir ünitede tartışılan araştırma bulguları ile ilişkili olan soruların kullanılması amacıyla, bir çok Düşündürücü Soru değiştirilmiştir. Çok beğendikleri bir soru eksik çıkarsa, öğretim elemanları onu Öğretim Elemanı El Kitabında bulabileceklerinden emin olsunlar.

Online Öğrenme Merkezi

Psikolojide Araştırma Yöntemleri'nin onuncu baskısının öğretim üyeleri ve Öğrenciler için hazırlanmış olan yardımcı kitapları www.mhhe.com/shaughnessy9e adresindedir. Shaughnessy, Zechmeister, ve Zechmeister'in, kitaptaki materyali desteklemek için hazırladıkları bu kaynakları, yazarlardan Jeanne Zechmeister bu onuncu baskı için güncelledi.

Öğrenciler İçin

Çoktan seçmeli, doğru-yanlış, ve eşleştirme soruları, problemler ve alıştırmalarla birlikte, çalışmaya yardımcı olarak kullanılabilir veya öğreniciye ev ödevi olarak verilebilir. Öğrenciler aynı zamanda öğrenme konularına, bir sözlüğe, ve her bölüm için online kaynaklara erişebilirler.

Öğretim Üyeleri İçin

Aşağıdaki kaynaklar, Psikolojide Araştırma Yöntemleri kitabını kullanan öğretim üyeleri içindir. İnternetten öğretim üyesi el kitabına erişim için gerekli parolayı almak için yerel McGraw-Hill satış temcileri ile iletişime geçin.

Psikolojide Araştırma Yöntemleri Öğretim Üyesi El Kitabı Güncel el kitabında bölümlerin özetleri ve konuları, bölümlerin gözden geçirme soruları ve cevapları, düşündürücü sorular ve cevapları, sınıfta tartışılacak meseleler ve problemler, araştırmanın eleştirel olarak nasıl okunacağına dair etkinlikler, öğrenciler için çalışma sayfaları, sınıf ve ev projeleri, öğretim üyesi için yardımcı ders ve tartışma materyali, ve PowerPoint slaytlarda veya çalışmalarda kullanılabilecek sayfalar bulunmaktadır.

PowerPoint Sunumları Her bölüme ait PowerPoint slaytları, bölümün ana konularını özetlemektedir.

Test Bankaları Her bölümdeki test bankalarında öğrencilerin bilgilerini test etmek için kısa cevaplar ve çoktan seçmeli sorular ve cevaplar yer almaktadır. Her bir soru, o sorunun değerlendirdiği husus (olgusal bir anlayışı mı, kavramsal bir anlayışı mı, yoksa metodolojik kavramların uygulanmasını mı değerlendirdiği) ile uyumludur. Test bankasına bilgisayardan da (EZ Test) ulaşılabilir. EZ test, sınavların kolayca hazırlanmasına olanak veren güçlü bir programdır. Online test hazırlamada güvenliği sağlamak için, EZ Test programında hazırlanan sınavlar, WebCT, Blackboard, ve EZ Test Online'a gönderilebilir. EZ Test, Quick Start Guide ile birliktedir; bir defa program kurulduktan sonra kullanıcılar, Kullanıcının El Kitabına ve öğreticiye erişebilirler. www.mhhe.com/eztest sayfasından daha fazla bilgiye erişilebilir.

COURSESMART

Bu kitap, eTextbook olarak www.coursesmart.com adresinde mevcuttur. Öğrencileriniz bu sitede basılı kitabı indirimli alabilir, çevreye etkilerini azaltabilir, ve öğrenme için güçlü bir Web aracına erişebilirler. CourseSmart eTextbook online olarak görülebilir veya bilgisayara indirilebilir. e-ders kitapları öğrencilere metin içinde arama yapma, metnin istenilen kısımlarını işaretleme, istenilen yerlere not ekleme ve sınıf arkadaşlarıyla not paylaşma olanağı sağlamaktadır.

CREATE ile dersinizi kendiniz şekillendirin!

Öğretim kaynaklarınızı McGraw-Hill'in size özel yayın programı doğrultusunda şekillendirin.

Etkin Olmak İçin:

- İstediğiniz içeriği bulun ve seçin, veya kendi içeriğinizi ekleyin.
- Kendi içeriğinizi dersi verme şeklinize göre düzenleyin. Ders materyalinizi mümkün olduğunca sık düzenleyin ve güncelleyin.
- Bir kapak ve öğrencileriniz için bir format seçerek, kitabınızı kişiselleştirin – basılı da olabilir bir eKitap da. Kitabınızın PDF kopyasını dakikalar içinde, veya basılı kopyasını birkaç gün içinde gözden geçirin.

TEŞEKKÜR

Kitabımızın bu 10'uncu baskısına birçok insanın toplamda yapmış oldukları katkılara hakkıyla teşekkür etmek neredeyse olanaksızdır. Eğer kitabı okuyan ve önerilerde bulunan şu araştırmacıların değişiklik önerilerinin hepsini yerine getirememişsek bundan üzüntü duyduğumuzu bildiriyor, kendilerine teşekkür ediyoruz: Jane Ashby (Central Michigan University), Austin Baldwin (Southern Methodist University), Nida Bikmen (Denison University), Nicole McCray (The University of Montana) ve Adriane Seiffert (Vanderbilt University). Ayrıca, Liz Zechmeister, Andy Apodaca ve Emil Posavac'a, çeşitli araştırma konularında faydalı bilgiler sağladıkları için teşekkür ediyoruz.

John J. Shaughnessy
Eugene B. Zechmeister
Jeanne S. Zechmeister

Türkçe Çeviriye Ön Söz

Bilimsel araştırma yöntemleri, tüm bilimlerin ortak tek alanıdır. Her bilim alanında elde edilen bulguların değeri, kullanılan yöntemin nesnelliğine ve geçerliğine bağlıdır. Bu nedenle, araştırma yöntemlerini öğretmek için yayımlanmış olan kitapların büyük çoğunluğu, farklı bilim alanlarında kullanılmak üzere genel bir içerikle hazırlanmıştır; örneğin, sosyal bilimlerin tamamına yönelik araştırma yöntemleri kitapları böyledir. Bu gibi genel içerikli kitapların önemli bir dezavantajı var: Yöntemler açıklanırken sık sık başvurulan (ve öyle olması gereken) araştırma örneklerinin de farklı alanlardan olması kaçınılmaz olmaktadır. Bu nedenle öğrenci, açıklanmakta olan yöntem ile örnek araştırma arasında ilişki kurmakta zorluk çekebilmektedir.

Elinizdeki kitap bu açıdan özgün bir kitaptır çünkü bilimsel araştırma yöntemlerinin sadece psikoloji alanında ne şekilde kullanılması gerektiğini açıklarken, yine sadece psikoloji alanında yapılmış araştırma örneklerini kullanmıştır.

Yazarların gıpta edilecek bir derecede bu kitaba zaman ve emek harcadıkları kolayca fark edilecektir. Psikoloji alanında araştırma yöntemleri kitabı yazmayı, adeta uzmanlık alanları hâline getirmişlerdir; yeni araştırmaları yakından izleyip her yeni baskıda yöntemlerin açıklanmasında en etkili şekilde kullanmışlardır. Psikolojinin çeşitli alt alanlarında hem "klasik" olmuş hem de yeni araştırma örnekleri, kendi başlarına yöntemi kavratabilecek kadar isabetli seçilmiş ve etkili kullanılmıştır.

Bu çevirinin en önemli sorunu, kolayca tahmin edilebileceği gibi, çoğu çevirilerde yaşanan terminoloji sorunu oldu. Türkçede farklı terimlerle ifade edilenler, hem İngilizceleri ile birlikte verildi hem de, bilerek, dilimizdeki bu farklı karşılıkları ile kullanıldı; zamanla bazılarının baskınlık kazanacakları düşüncesi, standartlık endişesinin önünde tutuldu.

Son olarak, çeviri ekibi bu kitabın "çeviri" izlenimi vermemesi için özel bir gayret gösterdi. Çoğu çevirilerde çeviri hatası olmadığı hâlde, okuyucu kendisine "Acaba ne demek istiyor?" sorusunu sorar. Elinizdeki çeviride, bu sorunun en aza indirilmesi hedeflendi.

Çeviri ekibi olarak, *Psikolojide Araştırma Yöntemleri* kitabının psikoloji öğrencilerimize yararlı olmasını diliyoruz.

Çeviri Ekibi (Alfabetik sıralı)

Prof. Dr. Sonia AMADO
Doç. Dr. Handan CAN
Doç Dr. İlyas GÖZ
Yrd. Doç. Dr. Osman İYİLİKÇİ
Yrd. Doç. Dr. Sema KARAKELLE

KISIM BİR

Genel Meseleler

BÖLÜM BİR

Giriş

BÖLÜMÜN ANA HATLARI

PSİKOLOJİ BİLİMİ
BİLİMİN BAĞLAMI
Tarihsel Bağlam
Sosyal ve Kültürel Bağlam
Ahlâkî Bağlam
BİR ARAŞTIRMACI GİBİ DÜŞÜNMEK
Medyada Yayımlanan Araştırma Bulgularını Değerlendirme
Araştırmaya Başlama
ÖZET

PSİKOLOJİ BİLİMİ

- Psikologlar, davranışa ve zihinsel süreçlere ilişkin soruları cevaplamak için teoriler geliştirirler ve araştırmalar yaparlar; verilen cevaplar, bireyleri ve toplumu etkileyebilir.
- Bir bilgi edinme aracı olan bilimsel yöntem, soruların sorulması ve cevapların alınması için mantığın ve yöntemlerin kullanılma usulleri demektir.
- Bilimsel yöntemin iki önemli özelliği, görgül bir yaklaşım ve kuşkucu bir tutumdur.

Sizin hem medyada hem de derslerinizde psikoloji alanında yapılmış birçok araştırma bulgusunu okuduğunuzu varsaymak, çok da tedbirsizlik sayılmaz. Eğer ders kitabınızın yazarlarına benziyorsanız, zihin ve davranış hakkında çok meraklısınız demektir. İnsanların (ve hayvanların) davranışları hakkında düşünmeyi seviyorsunuz. İnsanlar hakkında meraklısınız –neden öyle davranıyorlar, nasıl öyle insanlar oldular ve gelişim ve değişimleri ne şekilde sürecek. Ve kendi davranışınızı ve zihninizin nasıl çalıştığını merak edebilirsiniz. Bu gibi düşünceler, sizi diğer insanlardan ayırır –herkes zihni merak etmez ve herkes davranışın nedenleri üzerinde düşünmez. Ama eğer siz meraklıysanız, eğer insanların ve hayvanların neden öyle davrandıklarını merak ediyorsanız, psikolojide araştırma yöntemlerinin şaşırtıcı, heyecan verici ve evet, bazen zorlu yolculuğuna adım atmış sayılırsınız.

Birçok öğrenci psikoloji alanına, insanların hayatını iyileştirmeye ilgi duyduğu için girer. Ama hangi yöntemler ve müdahaleler insanlara yardımcı olur? Örneğin, psikoterapist olmayı amaçlayan öğrencilerin, uyumsuz davranış şemalarını teşhis etmeyi ve faydalı olan ve olmayan psikolojik müdahaleleri ayırt etmeyi öğrenmeleri gerekir. Psikologlar, davranış hakkındaki sorularını cevaplamak için teoriler geliştirerek ve araştırmalar yaparak, insanların hayatını iyileştirme konusunda anlayış ve kavrayış kazanırlar.

Çoğu psikolog, doğrudan insanların günlük hayatlarıyla ilgili konuları araştırır. Bir bilimsel dergi, *Psychological Science in the Puplic Interest*, sadece, hemen herkesin ilgisini çeken meselelere ilişkin davranış araştırmalarını yayımlamaktadır. Örneğin, 2012 yılına ait bir araştırma, internet üzerinden eş (partner) bulma ile geleneksel eş bulma yöntemlerini karşılaştırmaktadır (Finkel, Eastwick, Karney, Reis ve Sprecher, 2012). Bu araştırmacıların gözlemlerine göre, internet daha fazla potansiyel eş sağlıyor ama bu eşler büyük olasılıkla "mal" gibi görülebildikleri için, arayan kişi kendisini bunlardan birisine adamaya istekli görünmüyor. Bu dergideki bir başka makalede, öğrencilerin daha iyi öğrenebilmeleri için geliştirilmiş olan 10 teknik değerlendirilmiştir (Dunlosky ve ark., 2013). Bu araştırmacılar en iyi iki araştırma tekniğinin, çalışmaları zamana yayma ve alıştırma testleri alma teknikleri olduğu sonucuna varmışlardır. Bunu akılda tutarak, elinizdeki kitabın alıştırma testlerini Online Öğrenme Merkezi'nde bulabilirsiniz: www.mhhe.com/shaugnessy10e.

Psikologların en çok araştırmış oldukları konular arasında yer alan çok önemli bir araştırma sorusunu ele alalım: Medyadaki şiddetin etkisi nedir? Elli yıldan fazla bir araştırma döneminde yayımlanan yüzlerce araştırmada televizyon, film, video oyunları, internet ve müzikteki şiddetin etkisi araştırılmıştır; şimdi, bu araştırmalara dayanarak psikologlar, medyadaki şiddetin davranış, duygu, ve sosyal etkisi konusunda ne söylüyorlar? Yine yukarıdaki dergide

yayımlanan bir çalışmada Anderson ve ark., (2003), bazı kilit bulgulardan söz etmektedir:

—Medyada şiddete maruz kalmak, kısa ve uzun vadede, saldırganlık ve şiddet içerikli düşünce, duygu, ve davranış olasılığını arttır.
—Medyadaki şiddetin etkileri, çeşitli araştırma ve yöntemlerde, medya şekillerinde ve katılımcılarda tutarlıdır.
—Yakınlarda yapılan uzun süreli (boylamsal) araştırmalar, çocuklukta medya şiddetine maruz kalma ile yetişkinlikteki saldırganlık arasında bağlantı olduğunu düşündürmektedir; yetişkinlikteki bu saldırganlığa fiziksel saldırı ve eşlere uygulanan şiddet dâhildir.
—Araştırma bulguları psikologların, medyadaki şiddetin insanların saldırganlık bilişlerini ve fizyolojik uyarılmalarını "aktive ettiği" (kolaylaştırdığı) teorilerini desteklemektedir; bu kolaylaştırma, insanların gözlem yoluyla şiddet davranışlarını öğrenmeleri ve şiddete karşı duyarsızlaşmaları yoluyla meydana gelmektedir.
—Medyadaki şiddete karşılık olarak şiddete eğilim olasılığını etkileyen faktörler, seyircinin özellikleri (örneğin, seyircinin yaşı, kendisini saldırgan karakterle özdeşleştirme derecesi gibi) ve medya bağlamıdır (örneğin, şiddetin anlatımındaki gerçekçilik ve şiddetin sonuçları gibi).
—Medyadaki şiddetten herkes etkilenir.

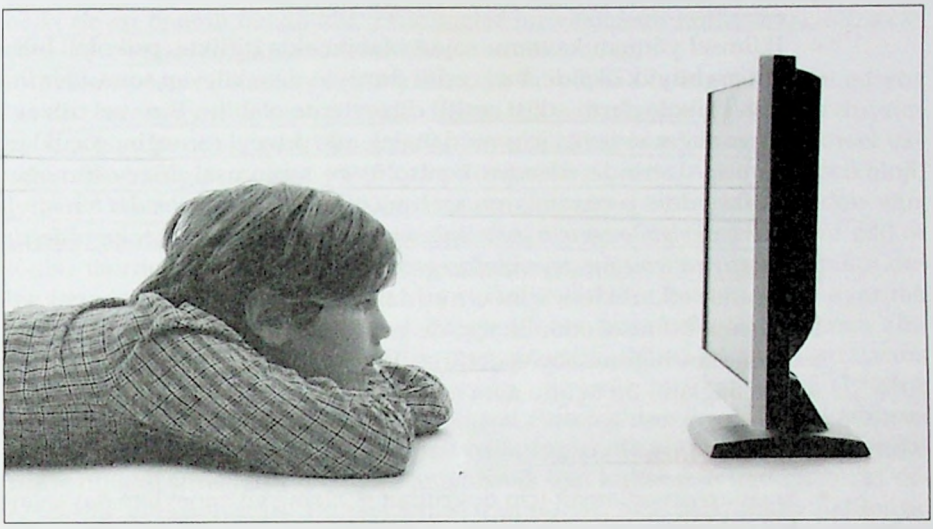
Birtakım araştırmalar gösteriyor ki çocuklar ve gençler, muhtemelen sadece uykudan daha az olmak üzere, haddinden fazla bir zamanı medya ile tüketiyorlar. Bu nedenle, yukarıda sıralanan araştırma bulgularının akla getirdiği şey şudur: Toplumdaki saldırganlık ve şiddetin yıkıcı sonuçlarını azaltmanın bir yolu, medyadaki şiddeti azaltmaktır. Gerçekten de psikolojik araştırmalar, ebeveynlerin televizyonda şiddet içerikli programları engelleyebilmelerini mümkün kılan Ş-anahtarının (Ş, şiddeti gösteriyor) geliştirilmesinde önemli rol oynadı (Anderson ve ark., 2003).

Geride daha çok araştırma sorusu duruyor. Önemli bir problem, şiddetin *pasif* seyircisi olmak ile (örneğin., televizyon) *aktif* olarak şiddet ile meşgul olmak (örneğin., video ve internet oyunları) arasında fark olup olmadığıdır (Şekil 1.1). Seyirci video oyunu oynarken aktif olarak şiddete katılırsa, medya şiddetinin etkisi daha fazla olur mu? Eğer aktif katılım saldırganlık eğilimlerini pasif seyircilikten daha fazla takviye ediyorsa, bu soruya "evet" cevabı verilebilir. Diğer araştırma soruları, şiddetin toplumumuzdaki etkilerini azaltmak için gerekli adımlar ile medyadaki şiddeti sınırlandırmanın özgür bir toplumda oynaması gereken rol üzerinedir. Belki de bu sorular ileride *sizin* araştırma sorularınız olacaktır veya belki de siz madde bağımlılığının nedenlerini veyahut ön yargının temellerini merak ediyorsunuz ve bunları keşfetmeye çalışacaksınız. Gerçekte, geride araştırılmayı bekleyen binlerce problem duruyor. Siz psikolojide araştırma konusunda çalışmaya devam ettikçe, bir gün psikologların insanın koşullarını iyileştirme çabalarına katkıda bulunabilirsiniz.

Temel Kavram

Psikologlar bilimsel yöntemi kullanarak davranış, düşünce ve duygular hakkındaki sorulara cevap ararlar. **Bilimsel yöntem** soyut bir kavramdır; soruların sorulmasını ve bu sorulara cevap vermek için kullanılan muhakeme ve yöntemleri açıklar. Psikologlar bilimsel yöntemi kullandıkları zaman *görgül bir*

ŞEKİL 1.1 Medyadaki şiddetin etkisi (a) pasif televizyon izleyici için, (b) aktif video oyuncusu için farklı mıdır?



(a)



(b)

yaklaşıma dayanırlar ve davranışın ve zihinsel süreçlerin açıklanmasında *kuş-kucu bir tutum* takınırlar. Biz bu iki özelliği, psikolojide araştırmaya girişin parçası olarak bu bölümde tartışacağız ve Bölüm 2'de ise bilimsel yöntemin başka özelliklerini ele alacağız.

BİLİMİN BAĞLAMI

- Bilim, en az üç bağlam içinde ortaya çıkar.

Bilimsel yöntem kavramı soyut olabilmekle birlikte, psikoloji biliminin uygulaması büyük ölçüde, bizi çeşitli düzeylerde etkileyen somut bir insan etkinliğidir. Psikologların etkisi çeşitli düzeylerde olabilir: Bireysel düzey (örneğin, saldırganlığın tedavisi için müdahale), aile düzeyi (örneğin, çocukların medya kullanımı üzerinde ebeveyn kontrolü) ve toplumsal düzey (örneğin, televizyondaki şiddet programlarını azaltma çabaları). *Böyle olmakla birlikte, bu etkilerinin verimli olabilmesi için psikologların, titizlikle planlanmış ve gerçekleştirilmiş olan bir araştırma temeline dayanmaları gerekir.*

İnsanın etkinlikleri, içinde meydana geldikleri bağlamdan çok etkilenir ve bu durum bilimsel etkinlik için de geçerlidir. Biz, bilimi çok etkileyen en az üç bağlamın olduğunu söyleyebiliriz: Tarihsel bağlam, sosyal-kültürel bağlam ve ahlâki bağlam. Bu üçünü aynı sırayla kısaca açıklayacağız.

Tarihsel Bağlam

- Soruları cevaplamak için doğrudan gözleme ve deneylere dayanan görgül bir yaklaşım, psikoloji biliminin gelişmesi için hayatî bir önem taşır.
- Bilgisayar devrimi davranışçı psikolojiden, psikoloji araştırmalarında baskın yaklaşım olan bilişsel psikolojiye geçişte kilit bir rol oynadı.

Aslında biz, psikoloji biliminin tam olarak ne zaman bağımsız bir bilim hâline geldiğini bilmiyoruz. Psikoloji, adım adım ortaya çıktı; kökleri, Aristo'nun düşüncelerine, daha sonraki filozoflardan Descartes ve Locke'un yazılarına ve nihayet 19'uncu yüzyıl fizyolog ve fizikçilerine kadar eskiye dayanır. Psikolojinin resmî başlangıcı, Wilhem Wundt'un Leipzig'de (Almanya) bir psikoloji laboratuvarı kurduğu tarih olarak kabul edilir (1879).

İlk psikologların 19'uncu yüzyılın sonlarında karşı karşıya kaldıkları kararlardan birisi, psikolojinin fiziksel bilimlerle daha yakın bir ilişki içine girmesi mi, yoksa felsefenin bir alt disiplini olarak kalması mı gerektiği meselesiydi (Sokal, 1992). Sinir sistemindeki iletişimi anlamak için psikofizik yöntemlerin ve tepki zamanı yöntemlerinin gelişmesiyle psikologlar, eninde sonunda düşüncenin kendisini ölçebileceklerine inandılar (Coon, 1992). Bu güçlü gözlem yöntemleriyle psikoloji ölçülebilir, laboratuvara dayalı bir bilim hâline gelme yoluna girdi. Bilimsel psikologlar yaptıkları zihin araştırmalarının, daha köklü bilimler olan fizik, kimya ve astronomideki kadar önem kazanacağını umdular (Coon, 1992).

Psikoloji biliminin ortaya çıkmasının önündeki engellerden birisi, 20'nci yüzyıla girilirken tinselciliğe ve ruhsal olaylara duyulan güçlü ilgiydi (Coon, 1992). Halk "zihnin" bu konularını psikolojinin alanı içinde gördü ve gaipten haber verme, telepati ve ölülerle haberleşme gibi konulardaki sorularına bilimsel cevaplar aradı. Bununla birlikte çoğu psikolog, bu sözde bilimsel konuları psikolojiden ayırmak istedi. Psikolojiyi bir bilim olarak kurmak için psikologlar, insan davranışı hakkındaki anlayışı geliştirmenin bir aracı olarak görgülcülüğe

Temel Kavram

(amprisizm) sarıldılar. **Görgül yaklaşım**, soruları cevaplamak için doğrudan gözlemi ve deney yapmayı ön planda tutar. Bu yaklaşım, bilimsel yöntemin belki de en önemli özelliğidir. Psikologlar bu yaklaşımı kullanarak, *doğrudan gözlenebilen* davranışlara ve deneyimlere odaklandılar.

Psikoloji her ne kadar görgül yaklaşımı ön planda tutmaya devam ediyor ise de başlangıcından bu yana çok değişti. İlk psikologlar esas olarak duyum ve algı konusundaki soruları cevaplamaya odaklandılar –örneğin, görsel yanılsamalar, imgelem gibi. Yirminci yüzyılın başlarında ABD'deki psikoloji, John B. Watson'ın başını çektiği davranışçı (behaviorist) yaklaşımdan ağır şekilde etkilendi. Psikolojik teoriler öğrenme konusuna odaklandı ve psikologlar teorilerini test etmek için çoğunlukla hayvanlar üzerinde yapılan deneylere dayandı. Davranışçılıkta, dışarıdan verilen bir uyarım ve bu uyarıma verilen gözlenebilir bir davranışsal tepki, "kara kutu" olarak nitelenen zihnin iç işleyişinden daha önemlidir. Davranışçılık, 20'nci yüzyılın ortalarına kadar psikolojide hâkim bakış açısı oldu. Bununla birlikte psikoloji, Ulric Neisser'ın1967 yılında yayımladığı *Bilişsel Psikoloji*'den önce, tekrar zihinsel süreçlerle ilgilenmeye başladı. Bilişsel psikologlar, ilk psikoloji laboratuvarlarında bilişsel süreçlerin niteliğini araştırmak için kullanılan tepki zamanı deneylerine geri döndüler. Bilişsel bakış açısı bugün hâlâ psikolojide hâkimdir, ve araştırmacıların zihnin biyolojisini araştırdığı bu dönemde biliş, sinirbilimin önemli konulardan biridir. Yirmi birinci yüzyılda, bilimsel psikolojinin gelişimi için büyük bir potansiyel vardır.

Bilişsel psikolojinin yükselmesinde önemli rolü olan bir etken, bilgisayar devrimidir (Robins, Gosling ve Craik, 1999). Bilgisayarın gelmesiyle davranışçılığın "kara kutu"su, bilgisayar benzetmesiyle temsil edildi. Psikologlar giriş (uyarım) ve çıkış (tepki) arasında bilgi işleme, depolama ve bulup çıkarma terimlerini kullandılar. Bilgisayar bilişsel süreçleri anlamak için yararlı bir benzetme olanağı sağlarken, kolayca ulaşılabilen güçlü bilgisayarlar bilişsel süreçlerin daha fazla alanını daha hassas şekilde ölçme olanağı sağladı. Bugün ABD'nin ve dünyanın her yerindeki psikoloji laboratuvarlarında, insanların düşüncelerini, duygularını ve davranışlarını ölçmek için kâğıt-kalem yerine artık bilgisayar kullanılmaktadır. Aynı şekilde, beyin görüntüleme (örneğin, fMRI) teknolojisindeki sürekli gelişmeler, psikoloji, biyoloji, kimya ve tıp alanlarında önemli bir disiplin olan nörobilimi geliştirecektir.

Psikolojinin tarihsel gelişiminde davranışçılıktan bilişsel nörobilime uzanan bu ana eğilimler, 20'nci yüzyılda psikolojide yaşananların "büyük resmi"ni temsil etmektedir. Bununla birlikte, daha yakından bakıldığında psikoloji biliminde araştırılan konuların çeşitliliği ortaya çıkar. Günümüz psikologları klinik, sosyal, organizasyonel, danışmanlık, fizyolojik, bilişsel, eğitim, gelişim ve sağlık psikolojisi gibi genel alanlarda araştırma yapmaktadırlar. Tüm bu alanlardaki araştırmalar, davranışın ve zihinsel süreçlerin karmaşıklığını anlamamıza yardım etmektedir.

Olağanüstü insanların parlak düşünceleri genellikle bilimi –ve özel olarak da psikolojiyi – değiştirdi. Galile'nin, Darwin'in ve Einstein'ın düşünceleri sadece bilim insanlarının kendi alanlarına bakış açılarını değil fakat aynı zamanda insanların kendilerini ve kendi dünyalarını anlayış biçimini de değiştirdi.

KUTU 1.1

PSİKOLOJİ VE NOBEL ÖDÜLÜ

İsveç Kraliyet Bilimler Akademisi her yıl çeşitli alanlarda yapılan araştırmalar için araştırmacılara ünlü Nobel Ödülünü vermektedir. Kasım 2002'de Daniel Kahneman, psikoloji alanında doktora yapmış ilk bilim insanı olarak bu ödülü kazandı. Kendisine bu ödül belirsizlik koşullarındaki sezgisel karar, insan muhakemesi ve karar verme konularında yaptığı çalışmaları nedeniyle verildi. Uzun yıllar birlikte çalıştığı meslektaşı Amos Tversky (1937-1996) ile yaptığı araştırmaların, ekonomi teorileri üzerindeki önemli etkileri nedeniyle ödüllendirildi (Kahneman, 2003). Kahneman Nobel Ekonomi Ödülünü, ekonomist Vernon Smith ile paylaştı; Vernon Smith iktisat alanında (bu kitapta önemli bir konu olan) laboratuvar deneyleri tasarlama konusundaki çalışmalarıyla bilinir.

Eğitimleri psikoloji alanında olmamakla birlikte, birçok bilim insanı doğrudan davranış bilimleriyle ilişkili araştırmaları için Nobel Ödülü kazandı (Chernoff, 2002; Pickren, 2003). Bunlardan bazıları aşağıdadır:

- **1904,** Fizyoloji veya Tıp: Ivan Pavlov, sindirim üzerine yaptığı araştırmalarla Nobel Ödülü kazandı; bu araştırmaları, sonradan klasik koşullama üzerine yaptığı çalışmalarını etkilemiştir.
- **1961,** Fizyoloji veya Tıp: Fizikçi Georg von Békésy, psiko-akustik (ses algısı) konusundaki araştırmalarıyla Nobel Ödülünü aldı.
- **1973,** Fizyoloji veya Tıp: Üç etolog, Karl von Frisch, Konrad Lorenz ve Nikolaas Tinbergen, ilk kez tamamen davranış üzerine olan araştırmaları için Nobel Ödülü kazandılar. Etoloji biyolojinin, canlının davranışlarını doğal ortamlarında araştıran alt alanıdır (bakınız, Bölüm 4).
- **1978,** Ekonomi: Herbert A. Simon, organizasyonel karar verme üzerine yaptığı çığır açıcı çalışmalarıyla Nobel Ödülü aldı (MacCoun, 2002). Kahneman, 2002'de aldığı Nobel Ödülünde Simon'ın araştırmalarının etkili olduğunu belirtti.
- **1981,** Fizyoloji veya Tıp: Nobel ödülü, ayrık-beyin işlemini kullanarak iki beyin yarım küresinin farklı işlevleri olduğunu gösteren Roger W. Sperry'ye verildi.



Bu bilim insanlarının ve başka birçoklarının başarıları, bilimde davranış araştırmalarının genişliğini ve önemini göstermektedir. Her ne kadar "Psikoloji Nobel Ödülü" yok ise de (ki matematikte de yoktur) bilim insanlarının çeşitli alanlarda yaptıkları araştırmaların bizim davranışı anlamamıza katkıda bulunduğu kabul edilmektedir.

Aynı şekilde, birçok olağanüstü insan, Nobel Ödülü sahipleri dâhil (bakınız, Kutu 1.1), psikolojinin gelişimini etkiledi (Haggbloom ve ark., 2002). Amerikan psikolojisinin ilk zamanlarında William James (1842-1910) zihinsel süreçleri araştırmak için kendi içebakış yöntemini geliştirdi, ve Sigmund Freud (1856-1939) kendi serbest çağrışım yöntemini kullanarak kişiliği, zihinsel bozuklukları, ve bilinçaltını anlamaya odaklandı. Davranışçılığın şöhreti artarken, B. F.

Skinner (1904-1990) davranışın deneysel analizine yöneldi. Başka birçokları gelişimsel psikoloji, klinik psikoloji, sosyal psikoloji ve bilişsel psikoloji gibi psikolojinin özel alanlarındaki düşünceyi büyük ölçüde etkiledi. Ümit ediyoruz ki sizler en çok ilgilendiğiniz alanlarda, geçmişte ve günümüzde etkili olmuş bu önemli psikologlar hakkında daha çok bilgi edineceksiniz.

Bilim aynı zamanda, birçok bilim insanının küçük katkılarıyla yavaş bir değişim gösterir. Bu aşamalı değişimi betimlemenin bir yolu, psikologluk mesleğinin gelişimini betimlemektir. Amerikan Psikologlar Birliği (APA) 1892'de 31 üyeyle kuruldu. Birliğin kuruluşunun 100'üncü yıldönümü olan 1992'deki üye sayısı 72.000'in üzerindeydi. Ve bugün, psikolojinin 54 alt alanında APA'nın üye sayısı 137.000'den fazladır. Bir diğer profesyonel organizasyon, Psikolojik Bilimler Birliği (APS), 1988'de kuruldu; amacı, psikolojideki bilimsel meseleleri ön plana çıkarmaktır. Hem APA hem de APS, psikologların kendi alanlarındaki son gelişmelerden haberdar olmaları için düzenlenen yıllık toplantıları desteklemektedir. Her bir organizasyon aynı zamanda, en son araştırmalardan hem kendi üyelerini hem de genel olarak toplumu haberdar etmek için, bilimsel dergiler yayımlamaktadır.

Siz, oluşmaya devam etmekte olan psikoloji tarihinin bir parçası olabilirsiniz. Hem APA hem de APS, hem lisans hem de lisansüstü psikoloji öğrencilerine hem eğitim hem de araştırma olanakları sağlamaktadır. Daimî üye veya öğrenci üye olarak APA ve APS'ye katılmak için gerekli bilgilere, bu birliklerin web sitelerine başvurarak ulaşabilirsiniz:

(APA) www.apa.org
(APS) www.psychologicalscience.org

Her iki web sitesi de yeni yapılmış olan önemli araştırma bulguları hakkında haberler, psikoloji yayınları hakkında bilgi (başlıca bilimsel dergiler için nispeten ucuz öğrenci abonelikleri dahil) ve birçok psikoloji organizasyonuna ait bağlantılara dair bilgiler sağlamaktadır. Göz atabilirsiniz.

Sosyal ve Kültürel Bağlam

- Sosyal ve kültürel bağlam araştırmacıların hangi konuları araştıracaklarını, toplumun araştırma sonuçlarını kabulünü ve araştırmaların yapılacağı yerleri etkiler.
- İnsanların başka kültürler hakkındaki görüşleri kendi kültürlerine ait anlayışlarından veya deneyimlerinden etkilendiği zaman, etnik merkezciyanlılık meydana gelir.

Bilim sadece kendi tarihsel bağlamından değil, aynı zamanda hâkim sosyal ve kültürel bağlamdan da etkilenir. Bu hâkim bağlam, bazen *hâkim fikir* veya *çağın ruhu* (zeitgeist) olarak da adlandırılır. Psikoloji araştırmaları ve uygulamaları, toplumla karşılıklı ilişki içinde var olur: Araştırma toplumu etkiler ve toplumdan etkilenir. Sosyal ve kültürel bağlam, araştırmacıların hangi konuyu araştıracaklarını, yapacakları araştırmayı destekleyecek mevcut kaynakları, ve ulaştıkları bulguların toplum tarafından kabulünü etkileyebilir. Örneğin, araştırmacılar kadın meseleleri ve etnik azınlıkların karşı karşıya oldukları meseleler üzerine yeni araştırma programları geliştirdiler. Bu durum kısmen, giderek daha fazla sayıda kadının ve azınlığın akademik kurumlarda temsil

edilmesinin bir sonucudur. Yayım hayatına 2013 yılında başlayan *The Journal of Latina/o Psychology*, göçmen, kültürel uyum ve kimlik ve giderek büyüyen Latin Amerika kökenli toplumun ihtiyaçlarını ele almaktadır. Sosyal ve kültürel tutumlar araştırmacıların sadece neyi araştırdıklarını değil fakat aynı zamanda, araştırmalarını ne şekilde yapacaklarını da etkileyebilir. Örneğin, toplumun iki dilliliğe karşı tutumu, araştırmacıların iki dil kullanmanın çocuklarda yol açtığı *problemleri* mi yoksa sağladığı *avantajları* mı ön planda tutacaklarını etkileyebilir.

Sosyal ve kültürel değerler insanların psikoloji araştırmalarının sonuçlarına nasıl tepki vereceklerini etkileyebilir. Örneğin, cinsel yönelim, çocukluktaki cinsel istismarın ortaya çıkarılması ve televizyonda şiddet gibi tartışmalı konularda yayımlanan araştırma sonuçları medyanın daha çok ilgisini çeker, çünkü kamuoyu bu konulara daha çok ilgi gösterir. Bu büyük ilgi bazen, araştırma bulgularının sosyal politikalar açısından nasıl yorumlanacağı ve nelerin yapılmasını gerektirdiği konusunda kamuoyu tartışmalarına yol açar. Çocukların cinsel istismarı konusunda *Psychological Bulletin*'da yayımlanan bir araştırmada (Rind, Tromovitch ve Bauserman, 1998) olduğu gibi, kamuoyunun tepkisi aşırı olabilir. Bu araştırmacılar çocukların cinsel istismarının etkisi üzerine yapılmış olan 59 araştırmayı inceleyip analiz ettiler ve şu sonuca vardılar: "Çocukta cinsel istismar, cinsiyet farkı gözetmeksizin, üniversite öğrencilerinde ağır ve yaygın bir zarara neden olmaz." (s. 46). Bu araştırma web'de çocuk istismarıyla mücadele sitesinde tanıtıldıktan sonra, "Dr. Laura" (talk şov programcısı Laura Schlessinger) araştırmacıların böyle bir niyeti *olmadığını* ama bu makalenin yetişkinlerin çocuklarla seks yapmalarını onayladığını belirterek APA'nın saygın bir dergisinde yayımlanmasını eleştirdi. ABD Temsilciler Meclisi 1999 yılında medyanın bu olumsuz ilgisine bir çözüm olarak, bu makalede bildirilen çalışmaya oy birliğiyle sansür kararı verdi. Aynı zamanda, *Psychological Bulletin*'da ve başka dergi ve kitaplarda yayımlanan eleştiriler ve karşı delillerle, tartışmalı bulgular üzerindeki bilimsel tartışma devam etmektedir (Dallam ve ark., 2001; Ondersma ve ark., 2001; Rind ve Tromovitch, 2007; Rind, Tromovitch, ve Bauserman, 2001). *American Psychologist* dergisinin bir sayısının tamamı, bu araştırmanın yol açtığı politik fırtınaya ayrıldı (Mart, 2002, Sayı, 3). Sağlam, görgül nitelikteki araştırma bulgularının bile bu şekilde kamuoyunda tartışılması, giderek güçlenen bir eğilim hâline geliyor. Araştırma bulgularına karşı olanlar, güçlü kişisel kanaatleri veya parasal çıkarları nedeniyle yasal, idarî ve siyasal saldırılarda bulunmaktadırlar (Loftus, 2003). Bu saldırıların, meşru bilimsel araştırma ve tartışmaları engelleyen talihsiz sonuçları olabilir.

Psikologların, çocukların cinsel istismarı gibi sosyal meselelere duyarlılığı, psikolojinin neden katı bir laboratuvar bilimi olarak gelişmediğinin nedenlerinden birisidir. Laboratuvara dayanan araştırmalar psikoloji araştırmalarının kalbi olmaya devam etmekle birlikte, psikologlar ve diğer davranış bilimciler okullarda, kliniklerde, iş yerlerinde, hastanelerde ve internet dâhil, laboratuvar dışında kalan diğer ortamlarda araştırmalar yapmaktadır. Gerçekte internet, psikologlar için giderek daha yararlı ve popüler bir araştırma aracı hâline geliyor (örneğin, Birnbaum, 2000). Tahminlere göre, 2012'nin sonlarına kadar tüm dünyada yaklaşık iki buçuk milyar internet kullanıcısı vardı; bunların yaklaşık yarısı Asya'da yaşıyordu (www.internetworldstats.com). Davranış bilimcilerinin planladıkları araştırmaları için şaşırtıcı büyüklükte ve çeşitlilikte bir "katılımcı havuzu"na erişmeleri, kısa zamanda mümkün olmaktadır (bk. örneğin, Birnbaum, 2000; Gosling, Vazire, Srivastava ve John, 2004; Skitka ve Sargis, 2005). Web, donanım

olarak bilgisayar ve katılımcı olarak insan kullanan hemen hemen her çeşit psikolojik araştırmaya olanak vermektedir (Krantz ve Dalal, 2000). Araştırmacıların, yapmayı planladıkları araştırma için katılımcı bulmalarının bir yolu, araştırma materyalini araştırma amaçlı çeşitli web sitelerine yollamaktır. Örneğin, APS'nin, internet kullanıcılarına psikoloji araştırmalarına katılma olanağı veren bir Web sayfası vardır. İnternette araştırma fırsatları hakkında bilgi edinmek için http://psyc.hanover,edu/research/exponnet.html sayfasını inceleyebilirsiniz. Psikolojideki özel araştırma yöntemlerini sizlere tanıttıkça, internette araştırma konusunda daha fazla bilgi vereceğiz. Bu araştırma biçiminin neden olduğu etik konular özellikle önemlidir (bk., Bölüm 3).

İnternete geniş erişim olanağı psikologlara, diğer ülkelerdeki araştırmacılarla ortak araştırma yapma ve psikoloji kavramlarını dünyanın her yanından katılımcılarla araştırma olanağı sağlamaktadır. Böyle olmakla birlikte, bir kısım psikoloji araştırması incelendiğinde görülmüştür ki APA'nın yayımladığı dergilerin önde gelen altı tanesinin katkıda bulunanları, örneklemleri ve editörleri, ağırlıklı şekilde Amerikalıdır (Arnett, 2008). Halbuki Amerikalılar dünya nüfusunun %5'inden daha azını temsil etmektedir ve dünyanın diğer yerlerindeki insanlar, Amerikalılardan çok daha farklı koşullarda yaşamaktadırlar. Şu hâlde birisi çıkıp, ağırlıklı olarak Amerikalıları araştıran bir psikoloji biliminin geçerli olup olmadığını sorabilir. Nitekim bazıları, psikoloji araştırmalarındaki katılımcıların çoğunun Batılı-Eğitimli-Sanayileşmiş-Zengin-Demokratik ülkelerden olduklarını ve bu durumun araştırma bulgularını çarpıttığını öne sürdüler bile (Jones, 2010).

Eğer biz bilimin sosyal ve kültürel faktörlerden etkilendiğini kabul ediyorsak, bir soru hâlâ cevap bekliyor: Kimin kültürü bilimi etkiliyor ve kimin kültürünün bilimi etkilemesi gerekiyor? *Kendi* kültürümüzün bakış açısını kullanarak *farklı* bir kültüre ait bireylerin davranışını anlamaya çalıştığımız zaman, potansiyel bir sorun ortaya çıkar (Şekil 1.2). Araştırmacılar kendi kültürlerine ait deneyimlerin ve değerlerin, onların diğer kültürlerde gözledikleri davranışların yorumlarını etkilediğini anlamadıkları zaman **etnik merkezcilik yanlılığı** ortaya çıkabilir. Amerikalıların katıldıkları psikoloji araştırmalarına ait bulgular diğer ülkelerdeki bireylerin davranış ve deneyimlerini yorumlamak için kullanıldığı ölçüde, potansiyel bir etnik merkezcilik yanlılığı mevcuttur.

Temel Kavram

Uzun yıllar boyunca psikologlar, kültürler arası psikoloji olarak adlandırılan araştırma alanında, kültürlerin benzerlik ve farklılıklarıyla ilgili meseleleri araştırdılar (Cohen, 2009). Bu alandaki araştırmaların çoğu, insanları, bireyci bir toplumda mı yoksa kolektivist bir toplumda mı ("bağımsız" – "karşılıklı bağımlı" da denilir) yaşadıklarına göre sınıflandırmaya odaklandı, ama göz önüne alınacak birçok kültürel boyut vardır; örneğin., cinsiyet, etnik köken, cinsel yönelim, din, sosyoekonomik statü ve insanların içinde yaşadıkları coğrafî bölge gibi (Cohen, 2009). Üstelik, insanlar çok kültürlüdürler, çünkü davranışları, düşünceleri ve değerleri bu boyutların hepsinden etkilenir.

Örneğin., psikolojik bir kavram olan *temsil* kavramını veya bir kimsenin kendi çevresini yönetmesini ve başkalarını etkilemesini ele alın. Araştırma bulgularına göre, sosyal statüsü yüksek olan bireyler temsile ve yönetmeye değer verirler, oysa çevrelerini etkileme olanağı kısıtlı olan düşük statüdekiler esnekliğe değer verirler (Snibbe ve Markus, 2005). Bu sosyokültürel farklılıklar,

ŞEKİL 1.2 Kendi kültürümüzün gözlüklerini kaldırarak; (a) yaşlılıkta güçlenme, (b) yetersizliklerden ziyade yeterlilikler ve (c) ev erkeği gibi konuları araştıran yeni araştırma düşüncelerine erişebiliriz.

(a)



(b)



(c)



bireylerin tercihlerini ve seçimlerini etkiler. "Kontrol altında"veya "bağımsız" gibi gözlemlenen eylemler iyiymiş gibi görüldükleri zaman, potansiyel bir etnik merkezci sapma görülür. Katrina kasırgasından kaçanlar, kalanlardan daha olumlu görüldükleri zaman durum böyleydi (Stephens ve ark., 2009). Kasırga esnasında kalanların, ki çoklukla düşük sosyoekonomik statüde insanlardı, kötü bir seçim yaptıklarına karar verildi. Bununla birlikte, bu insanlarla görüşüldüğü zaman kalma nedenleri olarak birbirlerine bağlılığı, gücü ve güveni öne çıkardılar. Araştırmacılar insanların davranışlarını etkileyen sosyal ve kültürel bağlamı gözden kaçırdıkları zaman etnik merkezci yanlılık ortaya çıkar. Kültürler arası araştırmalar, psikologların davranışı sadece bir hâkim kültürün bakış açısı ile araştırmaktan kaçınmalarına yardım eder ve araştırmayı yürütürken kültürel büyüteçler kullanmaktan kaçınmamız gerektiğini bize hatırlatır.

Ahlâkî Bağlam

- Araştırmanın ahlâki bağlamı, araştırmacıların etik davranış standartlarını en üst düzeyde tutmalarını talep eder.
- APA'nın etik yönetmeliği araştırmacılara rehberlik eder ve onların; araştırmada aldatma (deception) yapılırken ve hayvan kullanılırken karşılaşılan fayda/zarar ikilemlerini değerlendirmelerine yardım eder.

Bilim, bir "gerçeği arama" etkinliğidir. Bireysel olarak bilim insanı ve toplu bir girişim olarak bilim, içinde bilimsel etkinliğin meydana geldiği ahlâkî bağlamın, standartları en üst düzeyde karşıladığından emin olmak ister. Sahtekârlık, yalan ve gerçeğe aykırı açıklamanın, bir bilimsel etkinlikte rolünün olmaması gerekir. Ama bilim aynı zamanda insanî bir uğraştır ve sıklıkla doğruyu aramanın ötesinde kalan amaçları da içinde barındırır. Hem bilim insanları hem de onları çalıştıran bilim kuruluşları iş, para ve şöhret gibi ödüllerin olduğu bir oyunda rekabet eden oyunculardır. Örneğin, bir üniversitede çalışan bilim insanlarının yayımladıkları bilimsel araştırmaların sayısı çoğunlukla, terfi etme ve kadro alma gibi meslekte ilerleme kararlarını etkileyen başlıca etkendir. Bu koşullar altında, insanların düşündüklerinden daha fazla sayıda uygunsuz, ama görünüşe bakılırsa kaçınılmaz, bilimsel araştırma yapılmaktadır (Price, 2010).

Bilimsel dürüstlüğü bozan çeşitli etkinlikler vardır. Bunlardan bazıları veri uydurma, aşırmacılık (plagiarism), araştırma bulgularının bir kısmını seçerek yayımlama, araştırmaya katkıda bulunan insanlardan bahsetmeme, araştırma fonlarını kötüye kullanma ve insanlara ve hayvanlara etik dışı davranmadır (bk., Adler, 1991). Bazı ihlalleri tespit etmek, diğerlerinden daha kolaydır. Örneğin, verileri tamamen uydurma, bilimin normal işleyişinde, bağımsız araştırmacılar o uydurulmuş olan sonuçları bulamadıkları zaman veya bu uydurulmuş sonuçlar önceden yayımlanmış başka araştırmalarla uyumsuz olduğu zaman, tespit edilebilir. Ama sadece beklentileri karşılayan verileri yayımlama veya sonuçları yanıltıcı şekilde yayımlama gibi incelikli ihlalleri yakalamak zordur. Kasıtlı istismar ile kötü bilim arasındaki ayrım, her zaman belirgin değildir.

Araştırmacıları dürüst araştırma yapmaları için eğitmek ve mevcut birçok etik zorluğun üstesinden gelmelerine yardım etmek için, birçok bilim kuruluşu resmî etik yönetmeliğini kabul etti. Bölüm 3'te APA'nın insanlarla ve hayvanlarla yapılan araştırmalarda uyulması gereken etik prensipleri ele alacağız. Göreceğiniz gibi, etik dilemmalarla sık karşılaşılmaktadır. Örneğin, bir laboratuvar ortamında elinde silah olan erkek katılımcıların daha sonra daha saldırgan olup olmayacaklarını test eden araştırmayı ele alalım (Klinesmith, Kasser ve McAndrew, 2006). Araştırmacılar katılımcılara, ayrıntılara dikkat etmenin tat almaya duyarlılığı etkileyip etkilemediğini araştırdıklarını söylediler. Katılımcılar, iki dikkat koşulundan birisine yansız olarak atandılar. Bir grupta her bir katılımcı elinde bir silah tuttu ve silahın montajı ve sökülmesi üzerine bir yönerge yazdı. İkinci bir koşulda ise katılımcılar buna benzer yönergeler yazarken, Fare Kapanı oyunu oynadılar. Sonradan her bir katılımcının, içine bir damla acı sos konulmuş 85 gramlık bir su örneğinin tadına bakması istendi; görünüşe göre bu sosu suya, önceki katılımcılar koymuştu. Deneyin "tat duyarlılığı" bölümü buydu. Sonraki adımda katılımcılara su ve acı sos verildi ve kendilerinden sonraki katılımcılar için aynı şekilde bir soslu su örneği

hazırlamaları istendi. Suya ekledikleri acı sosun miktarı, saldırganlığın ölçütü (bağımlı değişken) olarak kabul edildi. Beklentiye uygun olarak, ellerinde silah tutan katılımcıların suya, silah tutmayanlara kıyasla, daha fazla acı sos koydukları bulundu; aynı sırayla, ortalama 13.61 gram ve 4.23 gram.

Bu araştırma birçok önemli soruya yol açtı: Hangi koşullar altında araştırmacının, deneyin gerçek amacı hakkında, katılımcıları aldatmasına izin verilmesi gerekir? Silahlar ve saldırganlık hakkında elde edilen bilginin yararı, aldatmanın neden olduğu risklerden daha fazla mıdır? Ellerinde silah tutan katılımcılar, araştırmanın amacının aslında silah ile saldırganlık arasındaki ilişki olduğunu bilseydi, suya daha az acı sos ekler miydi?[1]

Aldatma, araştırmacıların üstesinden gelmeleri gereken birçok etik meseleden sadece bir tanesidir. Bir diğer etik endişe konusunu, insan psikopatolojisini anlamak için bazen kullanılan hayvan katılımcıları düşünün. Bunun anlamı şudur ki bazen hayvanlar stresli ve hatta acı verici koşullara maruz bırakılabilir ve hatta ölüm sonrası inceleme amacıyla öldürülebilir. Hangi koşullarda psikoloji araştırmalarında hayvanlarla çalışmaya izin verilmelidir? Psikoloji araştırmalarının yol açtığı etik soruların listesi uzundur. Bu nedenle sizin, araştırmacılık hayatınızın henüz başındayken, APA'nın etik prensiplerini ve uygulamalarını öğrenmeniz büyük önem taşır; aynı şekilde, katılımcı, yardımcı veya araştırmacı olarak, sadece bilimsel uygunluk standartlarını en üst düzeyde karşılayan araştırmalara katılmanız da çok önemlidir. Ümit ediyoruz ki sizin bu araştırma yöntemleri eğitiminiz, sizin iyi araştırmalar yapmanıza ve hangi araştırmayı yapmanın daha iyi olduğunu ayırt etmenize olanak sağlayacaktır.

BİR ARAŞTIRMACI GİBİ DÜŞÜNMEK

- "Bir araştırmacı gibi düşünmek" davranışın ve zihinsel süreçlerin öne sürülen nedenleri hakkında, bu nedenler yayımlanmış bilimsel bulgulara dayansa bile, kuşkucu olmak demektir.
- Her ne kadar bilim insanları iddiaların daima olasılıklı olduğunu kabul ediyor iseler de, davranış hakkında öne sürülen bir iddianın en güçlü kanıtı, farklı araştırmaların buldukları uyumlu (converging) kanıtlardan gelir.

Bir psikoloji öğrencisinin atması gereken önemli bir adım, bir araştırmacı gibi düşünmeyi öğrenmektir. Bilimsel düşünmek demek, kişisel yanlılıkları ve peşin hükümleri bir tarafa koymak, ve en beklenmedik kanıtlarla karşılaşma cesaretini göstermek demektir (Lilienfeld, 2010). Bilim insanları her şeyden önce kuşkucudur. Davranışın ve zihinsel süreçlerin nedenlerine dair iddialara karşı kuşkucu bir tutum, psikolojide bilimsel yöntemin en önemli özelliğidir. Bilim insanı sadece "görmeden inanmayan" birisi değildir, aynı zamanda, aynı sonucu, muhtemelen kendi seçtiği koşullar altında, defalarca görmek isteyen kimsedir.

[1]Aldatma tekniğini kullanan herhangi bir araştırmanın önemli bir bölümü, deneyin sonundaki *deney sonrası bilgilendirme*'dir (debriefing). Bu bilgilendirme bölümünde, katılımcılara deneyin gerçek amacı açıklanır (bk., Bölüm 3). Klinesmith ve arkadaşlarının (2006) araştırmasındaki katılımcılara, asıl araştırma konusunun tat duyarlılığı değil, saldırganlık olduğu ve göstermiş oldukları herhangi bir saldırgan davranış için üzülmemeleri gerektiği söylendi. Katılımcıların hiç birisi, deney sonrası bilgilendirme esnasında deneyin gerçek amacından kuşku duyduğunu söylemedi. İlginç şekilde, Klinesmith ve arkadaşları, bazı katılımcıların, hazırlamış oldukları acı soslu suyun sonraki katılımcılara verilmemesine hayıflandıklarını bildirdi.

KUTU 1.2

KLİNİK PSİKOLOJİ VE BİLİM

Klinik psikologlar, hastalarını tedavi ederken psikoloji araştırmalarının en son bulgularını uygularlar mı?

Klinik psikolojinin uygulamalarına dair yeni bir eleştiride, saygın klinik psikologlar olan Dr. Timoty Baker, Richard McFall ve Varda Shoham bu soruya, yankı uyandıran bir "hayır" cevabı verdiler. Bu araştırmacıların klinik psikologların uygulamaları üzerine yaptıkları ve APA'nın *Psychological Science in the Puplic Interest* dergisinde Kasım 2008'de yayımlanan geniş analiz, *Newsweek* dâhil, medyada geniş şekilde yer aldı.

Geçen birkaç on yıldan fazla bir zaman klinik araştırmacılar, maliyet de göz önüne alındığı zaman, psikolojik tedavinin birçok zihinsel sorunun tedavisinde etkili olduğunu gösterdiler (örneğin, bilişsel davranışsal tedavi). Yine de, yukarıdaki araştırmacılara göre, nispeten az sayıda psikolog bu etkili tedavileri öğreniyor ve uyguluyor. Baker ve arkadaşlarına göre günümüzün klinik psikolojisi, 1800'lerdeki ve 1900'lerin başındaki bilimsel olmayan tıp uygulamalarına benziyor; o dönemin tıp doktorları kendi kişisel deneyimlerini, bilimsel uygulamalardan üstün tutuyorlardı. Araştırma gösteriyor ki günümüzün klinik psikoloğunun tutumu daha farklı değildir; tedavide neyin işe yarayacağı konusunda, görgül olarak desteklenmiş tedavilere ait bilimsel kanıtlardan ziyade, kendi kişisel düşüncelerini ön planda tutmaktadır. Gerçekte Baker ve arkadaşları ortalama klinik psikoloğun, görgül olarak desteklenmiş tedavilerle ilişkili araştırma bulgularının farkında olmadığını, ve muhtemelen, bir araştırmanın yöntemini ve bulgularını anlaması için gerekli bir bilimsel eğitiminin de olmadığını bildirmektedir.

Baker, McFall ve Shoham'a (2008) göre klinik psikologların eğitim programlarında acil değişiklikler yapılmalıdır; tıpkı tıbbı bilimsel bir temele oturtmak için 1900'lerin başında tıp eğitiminde köklü bir reformun yapılması gibi. Bilimsel bir temele dayanmadıkları sürece klinik psikologlar, günümüzdeki zihinsel ve davranışsal sağlık alanındaki önemlerini kaybetmeye devam edeceklerdir. Baker ve arkadaşları şuna inanmaktadırlar ki yüksek kaliteli, bilim odaklı öğrenim ve eğitim, klinik psikoloji eğitiminin temel özelliği olmalıdır, ve güçlü bir bilimsel temeli olmadığı sürece klinik psikoloji uygulamaları tatminkâr olmayacaktır.

Ümit ediyoruz ki bu kitabı kullanan öğrencilerin arasında klinik psikoloji ile ilgilenenler, çeşitli araştırma yöntemlerini öğrendikçe göreceklerdir ki bu araştırma yöntemlerine giriş kitabı, onların gelecekte başarılı ve etik bir meslek hayatlarının olabilmesi için bu alandaki ihtiyaçlarının sadece ilk adımıdır.

Araştırmacılar kendi öznel yargılarından çok, görgül kanıtlara dayanan sonuçlara ulaşmak isterler (bk., Kutu 1.2). En güçlü bilimsel kanıtlar, aynı araştırma sorusunu ele alan farklı araştırmalardan elde edilen birbiriyle uyumlu kanıtlardır. Davranış araştırmacıları kuşkucudurlar, çünkü davranışın karmaşık olduğunu ve psikolojik bir olayın çoğunlukla birçok faktörün etkileşimiyle ortaya çıktığını bilirler. Bu faktörleri keşfetmek, çoğunlukla zor bir iştir. Öne sürülen açıklamalar bazen çok zayıftır, çünkü bir olayı açıklayabilecek olan faktörlerin hepsi göz önüne alınmamıştır ve hatta farkına bile varılmamıştır. Davranış araştırmacıları aynı zamanda, bilimin bir insan çabası olduğunu da bilirler; insanlar hata yaparlar. İnsanın çıkarımları her zaman güvenilir olmaz. Bu nedenle bilim insanları, "yayımlanmış" bile olsa, "yeni keşifleri", işlemleri, ve olağanüstü iddiaları kuşkuyla karşılama eğilimindedirler.

Bilim insanlarının kuşkuculuğu onları, bilimsel eğitimi olmayan birçok insandan daha fazla temkinli olmaya yöneltir. Birçok insan, davranış hakkındaki sezgisel düşünceleri "olgu" olarak kabul etmeye eğilimlidir (Lilienfeld, 2010,

2012). Birçok yanlış düşünce, aksini gösteren psikolojik kanıtlara rağmen, doğru olarak kabul edilir; örneğin, romantik ilişkilerde zıtlıkların çekici olduğu, bastırılmış öfkeyi ifade etmenin öfkeyi azaltacağı ve dolunayda garip davranışların daha çok görüleceği düşüncesi böyledir. Çoktan seçmeli bir sınava girenlerin % 75'i, başka bir cevap doğru görünse bile, kendilerine doğru gelen ilk cevabı seçmeleri gerektiğine inanır (Kruger, Wirtz, Miller, 2005); siz bir öğrenci olarak böyle bir bulguyu ilginç bulabilirsiniz. Gerçekte, 70 yıldan daha uzun bir psikoloji araştırmaları literatürü gösteriyor ki cevaplarda yapılan değişmeler yanlıştan doğruyadır ve cevaplarını değiştirenlerin çoğu, puanlarını *yükseltmektedir*.

Elbette bilim insanları otomatik olarak, olaylar hakkında sağduyuya dayanan açıklamaların ve alışılmadık yorumların yanlış olacağını öne sürmezler. Onlar sadece, tüm iddiaların test edilmesine fırsat verilmesinde ve test edilemez nitelikte olanların reddedilmesinde ısrar ederler. Bilimsel kuşkuculuk, kolay inanan halkın, etkisi olmayan ilaç ve tedavileri pazarlayan sahtekarlara karşı, zengin olma düzenbazlıklarına karşı ve doğa olaylarının doğa üstü açıklamalarına karşı bir savunmasıdır. Bununla birlikte, aynı zamanda şu da akılda tutulmalıdır ki bilim insanının hayatında güvenin de kuşkuculuk kadar önemli bir rolü vardır. Bilim insanı kullandığı araçlara, katılımcılarına, meslektaşlarının araştırma bulgularına ve araştırmasını yaparken ulaştığı yargılara güven duyma ihtiyacı hisseder.

Size dedik ki bir bilim insanı gibi düşünmeniz için bulgulara ve iddialara karşı kuşkucu olmanız gerekir. Eğer bir suç ve mahkeme olayını anlatan bir kitap okuduysanız veya bazı popüler filmler ve televizyon oyunları izlediyseniz, kanıtlar ve iddialar hakkında zaten bir şeyler biliyorsunuz demektir. Dedektifler, avukatlar ve adlî alanda çalışan diğer meslek insanları, insanların davranışları hakkında iddialarda bulunabilmek için, çeşitli kaynaklardan birbiriyle uyumlu kanıtlar toplamaya çalışırlar. Bir miktar kanıt, bir insanın suçlu olduğundan şüphelenmek için yeterli olabilir ama bir insanı *mahkûm* edebilmek için çeşitli kaynaklardan sağlanan birbiriyle uyumlu kanıtlara gerek vardır. Psikoloji alanında çalışan bilim insanları da büyük ölçüde aynı şekilde çalışırlar –davranış ve psikolojik süreçler hakkında iddialarda bulunabilmek için kanıt toplarlar.

Elinizdeki kitabın ana amacı, farklı türden kanıtlara ve sonuçlara ulaştıran farklı araştırma yöntemlerini size detaylı olarak açıklamaktır. Bu ders boyunca ilerlemeye devam ettikçe göreceksiniz ki önemli ve farklı bilimsel prensipler vardır. Bu prensipler bir araştırmanın istatistiklerini veya bir davranış gözlemini rapor etmeye, faktörler (veya "değişkenler") arasındaki ilişkileri belirlemeye ve değişkenler arasında nedensel bir ilişki olduğunu ifade etmeye uygulanır. En güçlü bilimsel kanıtlar, bir duruşmada mahkûmiyet kararı vermek için gerekli olan birbiriyle uyumlu (converging) kanıtlara benzer. Araştırmacılar, bir deneyin tekrarlanmasına (replication) dayanarak, vardıkları sonuçlar için en güçlü kanıtlara sahip oldukları zaman bile durumları, bir kişiyi makul bir şüphenin ötesinde suçlu bulan jürilerin durumuna benzer. Hem araştırmacılar hem de jüriler gerçeği bulmak isterler ama vardıkları sonuçlar, nihayetinde bir olasılığa dayanmaktadır. Kesinlik, hem jürilerin hem de bilim insanlarının kavrayışlarını çoğunlukla aşar.

Bir araştırmacı gibi düşünmeyi öğrenerek, iki önemli beceriyi geliştirebilirsiniz. İlk beceri size, bilimsel bulguların daha verimli bir tüketicisi olma olanağını sağlayacaktır; böylece alanı daha iyi bilen ve daha isabetli kararlar verebilen birisi olacaksınız. İkinci beceri ise nasıl araştırma yapacağınızı öğrenmenize olanak sağlayacak; böylece, psikoloji biliminin gelişmesine katkıda bulunabileceksiniz. Kitap boyunca bilimsel yöntemin bu iki yönünü ayrıntılı şekilde anlatacağız ama

yine de bu bölümde kısaca özetleyeceğiz. İlk olarak, medyada yayımlanan araştırma iddialarını değerlendirirken neden bir araştırmacı gibi düşünmenin önemli olduğunu göstereceğiz. Daha sonra araştırmacıların, bilimsel yöntemi kullanarak kanıtlar toplamak istedikleri zaman nasıl işe başladıklarını betimleyeceğiz.

Medyada Yayımlanan Araştırma Bulgularını Değerlendirme

- Medyada yayımlanan bilimsel araştırmaların hepsi "iyi bilim" değildir. Okuduklarımızı ve işittiklerimizi sorgulamalıyız.
- Medya, orijinal araştırmanın özetlenmiş şeklini yayımlar; bu özet ise araştırmanın yöntemi, bulguları veya yorumuna ait önemli yönlerini vermeyebilir.

Psikoloji alanındaki araştırmacılar elde ettikleri sonuçları, basılı ve elektronik ortamda erişilebilen bilim dergilerinde yayımlarlar. Ama çoğu insan, bilimsel araştırmaları medyadan –internetten, gazetelerden, magazinlerden ve radyo ve televizyondan– öğrenir. Bu araştırmaların çoğu değerli araştırmalardır. Psikoloji araştırmaları çeşitli alanlarda insanlara yardımcı olur; örneğin, Alzheimer hastası yakınlarıyla iletişim kurmayı öğrenmek isteyen insanlar, tartışmaktan kaçınmayı, veya bağışlamayı öğrenmek isteyen insanlar gibi. Bununla birlikte, psikoloji bulguları medyada yayımlandığı zaman üç ciddi sorun ortaya çıkıyor. İlkin, medyada sunulan psikoloji, çoğunlukla herhangi bir araştırmaya dayanmıyor. Örneğin, her yıl yaklaşık 3500 kişisel gelişim kitabı yayımlanmaktadır, ama bunların ancak % 5'i bilimsel araştırmalara dayanmaktadır (Arkowitz ve Lilienfeld, 2006). Çoğu insan, kitapçıdaki rafları veya internette ilgili sayfaları incelerken iyi bir kişisel gelişim kitabını nasıl seçeceğini bilmez. İkinci olarak, medyada gerçek araştırma bulguları yayımlandığı zaman, araştırma "iyi" bir araştırma olmayabilir; yani, kabul edilmiş bilimsel işlemleri kullanarak yapılmış olmayabilir. İyi bir okuyucu, iyi araştırmayı kötüsünden ayırt etmek ister –hangilerinin sağlam bulgular olduğunu hangilerinin ise henüz doğrulanmadığını bilmek ister. Psikoloji bulgularının yayımlandığı her çeşit medyayı göz önüne alarak, araştırmaların çoğunun çok iyi olmadığını söylemek doğru olur. Bu nedenledir ki medyada okuduğumuz veya işittiğimiz araştırmaları sorgulamak için geçerli nedenlerimiz vardır.

Bilimsel araştırmalar medyada yayımlandığı zaman ortaya çıkabilen üçüncü problem, "çeviride" bir şeyler kaybolabilir problemidir. Medya yayınları, tipik olarak, orijinal araştırmaların özetleridir ve orijinal araştırmadaki yöntem, sonuçlar veya yorumun önemli kısımları bu özetlerde kaybolabilir. İşte bilimsel yöntemi ne kadar iyi öğrenirseniz, medyadaki bir araştırmanın kalitesini ve eksik yönlerini belirlemek için o kadar iyi sorular sorarsınız. Şimdi size birkaç soru tipi verebiliriz; bu soruları, medyada yayımlanmış olan bir araştırma örneğini inceledikten sonra sormak isteyebilirsiniz.

Mozart'ın müziğini dinlemenin, zekâyı arttıracağına inanıyor musunuz? Eğer inanıyorsanız, yalnız değilsiniz. Bir araştırmaya göre insanların %40'ı bu iddianın doğru olduğuna inanıyor (Chabris ve Simons, 2010). Dahası, daha siz bebekken veya hatta annenizin karnındayken, "uzmanlar" ebeveynlerinizi, size Mozart'ın müziğini dinletmeye teşvik etmiş olabilirler.

Tüm bunlar, bir psikoloji araştırmasının medyada heyecanla bildirilmesinin bir sonucudur ve "Mozart etkisi" adıyla bilinir. Müzik dinlemenin zekâyı arttıracağı düşüncesi, merak uyandıran ve şaşırtıcı bir fikirdir ve insanların "kolay çözüm" arayışlarının, kâr peşinde koşan medya ve şirketler tarafından kullanılmasının

yollarından birisidir (Schellenberg, 2005). Peki biz "Mozart etkisi" hakkında ne biliyoruz? Medyada bir psikoloji araştırması ile karşılaştığınız zaman atabileceğiniz iyi bir ilk adım, *araştırmanın yayımlandığı orijinal kaynağa gitmektir.*

Mozart etkisi makalesi ilk defa, saygın bilim dergisi *Nature*'da (Rauscher, Shaw ve Ky, 1993) yayımlandı. Basit bir deneyde üniversite öğrencileri ya Mozart'tan bir parça dinlediler ya sessizce oturdular ya da rahatlama yönergeleri dinlediler. Bu üç koşulun her birisi 10 dakika sürdü ve daha sonra bir mekânsal akıl yürütme testi aldılar. Mozart müziği dinleyen grubun bu testteki başarısı, diğer iki grubun başarısından daha yüksek oldu; ama bu fark, ilave bir 10-15 dakikalık aradan sonra ortadan kalktı.

Her ne kadar bu bulgular bir bilim dergisinde yayımlanabilecek kadar sağlam bulundu ise de bu bulguların uygulanması çok zayıftır. Milyonlarca ebeveyn, 36 üniversite öğrencisinde çok özel bir akıl yürütme testiyle gösterilmiş, ve en çok 15 dakika sürmüş olan bir etkiye dayanarak, bebeklerine "zekâ müzikleri" dinletmeye teşvik edilmişlerdir. Mozart etkisi üzerinde yapılan sonraki araştırmaların sonuçları karışık oldu. Bazı araştırmacılar orijinal sonuçları tekrarlayamadılar (replication) ve Mozart etkisi gözlendiği zaman, bu etkinin küçük olduğu bulundu (Chabris, 1999; Steel ve Bass, ve Crook, 1999). Dahası, araştırmacılar gösterdiler ki test başarısındaki artışın nedeni olumlu duygu durumu ve hoş bir uyarımı dinlemenin sonucu olarak artan uyarılmadır –bu, psikolojide iyi bilinen bir bulgudur (Thompson, Schellenberg, ve Husain, 2001).

Emin bir şekilde denilebilir ki Mozart etkisine olan yaygın inanışta, orijinal bulgunun çevirisinde "bir şey kaybolmuştur." Bu hikâyeden alınacak ders, medyada yayımlanan araştırma bulgularına karşı kuşkucu olmaktır. Mümkün olduğu zaman orijinal kaynağa gidin veya sağlam bilimsel kanıtlara dayanan düşünceler bulun. Sizin işiniz, okuduğunuz veya işittiğiniz şeyi sorgulamaktır; bizim işimiz ise araştırma bulgularını değerlendirmeniz ve anlamanız için size gerekli olan araştırma yöntemleri bilgisini sağlamaktır.

Araştırmaya Başlama

- Bir araştırmaya başlarken öğrenciler, bilim dergilerindeki, kitaplarındaki, ve derslerindeki konuları inceleyerek "Hangi konuyu araştırmalıyım?" sorusunu cevaplayabilirler.
- Bir araştırma hipotezi, bir olay hakkında yapılan geçici bir açıklamadır; hipotez çoğunlukla, hangi sonucun beklendiğini açıklayan bir kestirim şeklinde ifade edilir.
- Araştırmacılar hipotezleri çeşitli şekillerde kurarlar, ama daima, araştırmalarına başlamadan önce yayımlanmış psikoloji araştırmalarını incelerler.
- Araştırma sorusunun iyi bir soru olup olmadığına karar vermek için araştırmacılar sorunun bilimsel önemini, kapsamını, araştırmanın muhtemel sonuçlarını ve psikoloji biliminin ilerlemesine katkısının olup olmayacağını göz önüne alırlar.
- "Çok yöntemli yaklaşım" (araştırma sorularına cevap bulmak için çeşitli araştırma yöntemleri ve ölçümleri kullanan yaklaşım), psikolojinin davranış ve zihni anlaması için en ümit vaat eden yaklaşımıdır.

Sizler psikoloji alanında çalışan araştırmacıların ne şekilde kanıt topladıklarını öğrenmeye başladıkça, size çeşitli uzman araştırmacıların öğütlerini

aktaracağız ki bu, araştırmacılığın en temel yönlerinden birisidir. Bu bölümü, araştırmacıların bir araştırmaya başlarken kendilerine sordukları şu üç soru çerçevesinde düzenleyeceğiz:

—Neyi araştırmalıyım?
—Araştırmamda test edeceğim hipotezi nasıl kurabilirim?
—Araştırma sorum, iyi bir soru mudur?

Araştırma konusu seçileceği zaman, çoğu psikoloji öğrencisi dikkatini hemen psikopatolojiye ve zihinsel sağlıkla ilgili meselelere çevirir. Diğer öğrenciler ise bellek, problem çözme ve karar verme süreçleri gibi insan bilişinin çeşitli meselelerini merak ederler. Ve başkaları, gelişim psikolojisi ve sosyal psikoloji konuları ile ilgilenirler. Psikoloji, geniş bir araştırma olanağı sunan zengin bir alandır; nitekim, psikoloji alanındaki araştırmaları yayımlayan mübalağasız yüzlerce bilimsel derginin varlığı da bunu göstermektedir. Standart bir psikolojiye giriş kitabının "içindekiler" kısmını incelediğiniz zaman, psikolojinin birçok araştırma alanı hakkında çabucak bilgi edinebilirsiniz. Elbette üniversitenizde verilen örneğin, anormal psikolojisi, bilişsel psikoloji ve sosyal psikoloji gibi birçok derste daha spesifik bilgi bulabilirsiniz.

Psikolojide araştırma sorularıyla sadece öğrenciler ilgilenmez. Temmuz 2009'da *Perspectives in Psychological Science* dergisi bir sayısının tamamını, psikolojide araştırma sorularının ve psikolojinin geleceğinin tartışılmasına ayırmıştır (Diener, 2009). Psikolojinin çeşitli alanlarında çalışan ve önde gelen araştırmacıları, kendi alanlarındaki önemli soruları belirlediler –örneğin, zihin-beyin ilişkisini, evrimsel psikolojiyi ve hatta insan-robot etkileşimini ele alan sorular. Bir araştırma sorusuna cevap aranırken, bu makaleleri okumak iyi bir başlangıç noktası olabilir.

Öğrenciler ilk araştırma konularını çoğunlukla psikoloji öğretmenleriyle birlikte belirlerler. Öğretim üyelerinin çoğu araştırma yapar ve öğrencilerini bu araştırmaya dâhil etmeye isteklidir. Bu durumda araştırmaya katılmak için size düşen, sadece kendisine sormaktır. Psikoloji bölümlerinin ayrıca, öğrencilerinin araştırma düşünceleri geliştirmelerine yardımcı olmak için başka birçok kaynağı da vardır. Bunlardan birisi belirli bir konunun tartışıldığı seminerlerdir (colloquium). Böyle bir seminer, bir araştırmanın sunulduğu formel bir yerdir; bazen başka üniversitelerden de gelebilen araştırmacılar seminerde kendi görüşlerini ve araştırmalarını fakülte ve bölümdeki öğrencilere sunarlar. Sizin bölümünüzde yapılan seminer duyurularını izleyin ve katılın.

Bir araştırma konusunu nerede ve nasıl geliştirmeye başladığınızın hiçbir önemi yoktur; bir araştırmaya başlamanın önemli bir ilk adımı, yayımlanmış psikoloji araştırmalarını incelemektir. Psikoloji araştırmalarını internet kullanarak araştırma nispeten kolay ve hatta heyecanlı bir iştir. Bu kitabın on üçüncü bölümünde, bilgisayar kullanmak dâhil, psikoloji araştırmaları literatürünü nasıl inceleyeceğinizi özetleyeceğiz.

Bir araştırmaya başlamadan önce neden psikoloji literatürünü araştırmanız gerektiğinin çeşitli nedenleri var. Açık bir nedeni şudur ki sizin sorunuzun cevabı önceden zaten verilmiş olabilir. Bir başkası aynı soruyu ele almış ve en azından kısmen bir cevap vermiş olabilir. Büyük olasılıkla, sizin araştırma sorunuzla ilişkili araştırma bulguları ile karşılaşacaksınız. Belki siz kendi araştırma sorunuzun önceden araştırılmış olmasından biraz hayal kırıklığı yaşayacaksınız ama düşünün ki sizinkiyle aynı veya benzer konularda araştırma yapan insanlarla karşılaşmanız, sizin düşüncelerinizin değerli olduğunu gösteriyor.

Daha önce hangi konuların araştırıldığını ve cevaplandığını dikkatle incelemeden araştırma yapmak, ilginç veya eğlenceli (ve elbette kolay) olabilir; siz buna "hobi" diyebilirsiniz ama biz bu yaptığınıza bilim diyemeyiz. *Bilim, birikimli bir etkinliktir –yeni araştırmalar, önceki araştırmalara dayanır.*

Bir defa kendi düşüncenizle ilişkili olan önceki araştırmaları belirledikten sonra, bunları okudukça literatürdeki uyumsuzlukları ve çelişkileri keşfettiğinizi göreceksiniz. Ayrıca göreceksiniz ki okuduğunuz araştırmaların bulguları katılımcıların niteliği ile veya araştırmanın yapıldığı koşullar ile sınırlıdır veyahut/ test edilmeyi bekleyen psikoloji teorileri olduğunu göreceksiniz. Böyle bir keşif size, peşinden gidebileceğiniz iyi bir araştırma fikri verir.

Psikoloji araştırmalarını okuduğunuz ve muhtemel araştırma soruları üzerinde düşündüğünüz zaman, psikoloji araştırmalarına ait sonuçların toplumsal hayata nasıl uygulandıkları konusunda da düşünebilirsiniz. Psikolojide araştırmanın nasıl yapılacağını öğrendikçe, bu bilginin insan ve hayvan hayatını geliştirmek için ne şekilde kullanılacağını da ele alabilirsiniz.

Funder'e (2009) göre psikoloji en şanslı bilimdir, çünkü örneğin "İnsanlar neden öyle davranıyorlar?" gibi en ilginç sorular psikoloji alanındadır. Ama bazı insanlar, tüm bu ilginç soruların (ve cevaplarının) nerede olduğunu öğrenmek için psikoloji dergilerini tararlar ve kişilik ve sosyal psikoloji araştırmalarını eleştirirler, çünkü araştırılan "davranış" en çok bilgisayarda anketlere verilen cevaplardan veya tepki zamanlarından ibarettir (Baumeister, Vohs ve Funder, 2007). Buna karşılık olarak birçok psikolog, insanların gerçekten yaptıkları (örneğin, yeme, politik görüşler, flört gibi) ve "gerçek" davranışa daha çok benzeyen şeyleri betimleyen araştırmaların daha fazla yapılması için çağrıda bulundu (Funder, 2009; Rozin, 2009). Belki de bu, araştırmalarınızı planlarken aklınızda bulundurmanız gereken bir şeydir.

Nihayet, Sternberg'in (1997) işaret ettiği gibi, bir araştırma sorusu seçme işi hafife alınmamalıdır. Bazı soruları sormaya hiç değmez çünkü bunların cevabı, psikoloji biliminin ilerlemesine bir katkıda bulunmaz. Bu sorular, tek kelimeyle anlamsızdır veya en fazla önemsizdir. Sternberg (1997), psikoloji araştırmaları alanında yeni olan öğrencilerin, iyi bir araştırma sorusu bulduklarına karar vermeden önce, çeşitli sorular sormalarını öneriyor:

—Bu soru neden önemli olabilir?
—Bu sorunun kapsamı nedir?
—Bu araştırma projesini gerçekleştirirsem, muhtemel sonuçları nelerdir?
—Bu sorunun cevabı, psikoloji biliminin ilerlemesine ne kadar katkıda bulunacak?
—Neden birileri bu sorunun cevaplarıyla ilgilensin?

Araştırma sürecinin başında bu sorulara cevap bulma, araştırma danışmanlarının ve kendi başlarına başarılı araştırmalar yapmış olan başkalarının rehberliğini gerektirir. Ayrıca, ümit ediyoruz ki, psikolojideki teoriler ve araştırmalar hakkında daha fazla şey öğrendikçe ve bu kitapta betimlediğimiz birçok ilginç ve anlamlı psikoloji araştırmasını okudukça, sizin bu sorulara cevap verme yeteneğiniz artacaktır.

İLERİ ALIŞTIRMA

Bu alıştırmada, her bir sütundan bir madde seçerek hipotezler oluşturun. İlk sütundaki bir olay veya bir davranışı, ikinci sütundaki bir sonuçla, ve daha sonra üçüncü sütundaki mümkün bir açıklamayla birleştirin. Bu şekilde oluşturulan örnek bir hipotez, koyu yazılarak gösterilmiştir.

Olay veya Davranış	Sonuç	Açıklama
1 zayıf beden fotoğraflarını görme	1 yardımda artış	1 olayların tekrar yorumlanması
2 11 Eylül terör saldırısı	2 sağlık yardımları	2 daha çok empati
3 duygusal olaylar hakkında yazma	3 trafikte ölümlerin artışı	3 kendini ideal ile karşılaştırma
4 davranışları ve duruşları taklit etme	4 beden tatminsizliği	4 uçaktan korkma

Örnek Hipotez: Duygusal olaylar hakkında yazmak, muhtemelen yazıdaki olaylar tekrar yorumlanacağı için, sağlık yardımına neden olur [Pennebaker ve Francis, 1996].

Temel Kavram

Bir sonraki karar, biraz zor olabilir. Araştırmacılar işe koyuldukları zaman hipotezlerini belirlemeye çalışırlar. Bir **hipotez**, bir olayın geçici açıklamasıdır. Çoğunlukla bir hipotez, kestirimin bir açıklamasıyla birlikte, bazı sonuçlara dair bir kestirim şeklinde ifade edilir. Biz bu bölümde daha önce bir araştırma hipotezi önermiştik: Medyadaki şiddetin saldırganlığa etkisi, video oyunlarında, pasif televizyon izlemeye kıyasla daha kuvvetli olabilir, çünkü videoda oyuncular, saldırganlık eğilimlerini arttıran saldırgan eylemlerle aktif şekilde meşgul olmaktadırlar (Alternatif bir hipotez ise şunu iddia edebilir: Video oyunlarının etkisi daha *zayıf* olabilir, çünkü oyuncular saldırganlık dürtülerini serbest bırakıyorlar ki pasif televizyon izleyicisi bu olanaktan yoksundur.).

McGuire (1997), bilimsel olarak test edilebilecek bir hipotez kurmak için keşfe götürebilecek 49 basit kural ("heuristics") belirledi. Biz burada bu 49 kuralın hepsini ele alamayız ama bu kuralların bazılarını özetleyerek, McGuire'in ana fikri hakkında bilgi verebiliriz. Örneğin, O'na göre bir araştırma hipotezi

—genel bir eğilim veya prensipten sapmalar (gariplikler, istisnalar) üzerinde düşünerek;
—bir görevde veya spesifik bir problemle karşı karşıya kaldığımız zaman nasıl davranacağımızı hayal ederek;
—çözümleri bilinen benzer problemler üzerinde düşünerek;
—bir kimseyi veya olayı detaylı şekilde sürekli gözleyerek (örneğin, bir "vaka incelemesi");
—davranışa dair açık bir sonuçla uyuşmayan örnekler bularak;
—diğer bilim alanlarına ait düşünceleri veya teorileri kullanarak oluşturulabilir.

Ama açıktır ki bir araştırma sorusu ve hipotezi belirlemeniz, o araştırmayı nasıl yapacağınızı hemen söylemez. Anlamak istediğiniz şey tam olarak nedir? Bu soruyu cevaplamak, bu kitap boyunca ele alacağımız başka kararları da vermeniz gerektiği anlamına gelecektir. Bir araştırmacı olarak,

TABLO 1.1 ARAŞTIRMA SÜRECİNİN AŞAMALARI

Aşama	Nasıl?	Bölüm
Bir araştırma sorusu bulun.	• Etnik merkezci yanlılığa dikkat edin.	1
	• Araştırma yapmada kişisel deneyim kazanın.	1
	• Psikoloji literatürünü okuyun.	1, 13
Bir araştırma hipotezi kurun.	• Probleminizle ilgili teorileri okuyun.	1, 2
	• Kişisel deneyimi göz önüne alın, istisnaları düşünün ve önceki araştırmalardaki uyumsuzluklara dikkat edin.	1
İşevuruk tanımlar yapın.	• Aynı yapıları önceki araştırmaların nasıl tanımladıklarına bakın.	2
	• Araştıracağınız değişkenleri belirleyin.	2
Bir araştırma deseni seçin.	• Bir örneklem belirleyin.	4, 5
	• Araştırma sorunuzun bir şeyi betimlemeyi mi, kestirmeyi mi, yoksa bir neden sonuç ilişkisi bulmayı mı hedeflediğine karar verin.	2
	➤ Betimleme ve kestirim için, gözleme dayanan ve korelasyonel desenler seçin.	4, 5
	➤ Nedensel bir araştırma sorusu için deneysel bir desen seçin.	6, 7, 8
	➤ Eğer küçük bir grubu veya bir kişiyi anlamak ve tedavi etmek istiyorsanız vaka çalışması desenini seçin.	9
	➤ Eğer kontrol olanağının düşük olduğu bir ortamda nedensel bir ilişki araştırıyorsanız, yarı deneysel bir desen seçin.	10
Araştırmanızı etik açısından değerlendirin.	• Araştırmanın kendisinin ve katılımcıların sağlığının korunması için alınmış olan önlemlerin potansiyel risklerini ve yararlarını belirleyin.	3
	• Etik kuruluna bir öneri sunun.	3
	• Yetkililerin iznini alın.	3, 10
Verileri toplayın ve analiz edin; araştırmanın sonuçlarını belirleyin.	• Verileri tanımaya başlayın.	11
	• Verileri özetleyin.	11
	• Verilerin anlamını onaylayın.	12
Araştırma raporunu yazın.	• Bulguları bir psikoloji konferansında sunun.	13
	• Araştırmanın yazılı bir raporunu bir psikoloji dergisine sunun	13

kendinize "Nicel bir araştırma mı yoksa nitel bir araştırma mı yapmam gerekir? Araştırmak istediğim değişkenlerin niteliği nedir? Geçerli ve güvenilir davranış ölçümlerini nasıl bulabilirim? Araştırma soruma en uygun araştırma yöntemi hangisidir? Ne tip bir istatistiksel analize ihtiyacım olacak? Seçtiğim yöntemler ahlâkî ve etik standartları karşılayacak mı?" Bilimsel süreçle bağlantılı bu ve başka adımlar Tablo 1.1'de görülmektedir. Eğer bu sorulardaki ve Tablo 1.1'deki terimler size yabancı gelirse, hiç endişe etmeyin. Psikolojide araştırma yöntemleri kitabınızda ilerledikçe, araştırma sürecinin bu adımlarını da öğreneceksiniz. Tablo 1.1, kendi araştırmanıza başlayacağınız zaman sizin için iyi bir kılavuz olacaktır.

Bu kitap size, psikologların bilimsel yöntemi nasıl kullandıklarını açıklayacaktır. Bildiğiniz gibi psikoloji, çeşitli alt araştırma alanları ve cevap arayan birçok sorusu olan bir disiplindir. Bir tek araştırma yöntemi, psikologların davranış ve zihinsel

Temel Kavram

süreçlere ilişkin tüm sorularını cevaplayamaz. Bu nedenle, sorularımızı cevaplamamız için en iyi yaklaşım, **çok yöntemli yaklaşım**dır –yani, bir sorunun cevabını bulmak için çeşitli araştırma yöntemleri ve davranış ölçümleri kullanma. Bu kitabın amacı, "araç kutunuz"u bir araştırmayı yapabilmeniz için gerekli olan stratejilerle doldurmanıza yardım etmektir. Bu kitabı okudukça öğreneceğiniz gibi, hangisi olursa olsun bir tek davranış yöntemi veya ölçümü, araştırma sorusunu tam olarak cevaplama yeteneği açısından kusurlu veya noksan olabilir. Araştırmacılar çoklu yöntem kullandıkları zaman, bir yöntemin kusurunu diğer yöntemler telafi edebilir. Bu nedenle, çok yöntemli yaklaşımın önemli bir avantajı şudur ki araştırmacılar davranış ve zihinsel süreçler hakkında daha kapsamlı bir anlayış edinirler. Ümit ediyoruz ki bu kitapta betimlenen araştırma yöntemlerini kullanarak, psikoloji alanına ait sorularınızı cevaplayacak bir yola girmiş olacaksınız.

ÖZET

Psikologlar, teoriler geliştirerek ve araştırmalar yaparak davranışı ve zihinsel süreçleri anlamaya çalışırlar. Psikoloji araştırmaları, bireyler ve toplum üzerinde önemli bir etki yapabilir; bunun bir örneği, medyadaki şiddetin olumsuz etkilerini gösteren araştırmadır. Araştırmacılar, davranışı anlamak için görgül yaklaşımı önde tutan bilimsel yöntemi kullanırlar; görgül yaklaşım, soruları cevaplamak için doğrudan gözleme ve deneye dayanır. Bilimsel etkinlik tarihî, sosyal-kültürel ve ahlâkî bir bağlam içinde yapılır. Tarihsel olarak bilgisayar devrimi, psikolojinin ağırlıklı olarak davranışçı yaklaşımdan bilişsel yaklaşıma kaymasında etkili oldu. Birçok psikolog, geçmişte ve günümüzde, psikolojinin çeşitli alanlarının gelişimine katkıda bulundu.

Sosyal-kültürel bağlam psikoloji araştırmalarını, araştırmacıların hangi konuları araştıracakları ve toplumun o araştırma bulgularını kabulü açısından etkiler. Kültür, etnik merkezci yanlılık olduğu zaman da araştırmaları etkiler. Bu yanlılıkta araştırmacılar, farklı bir kültürde yaşayan insanların davranışlarını, kendi kültürlerinin çerçevesine veya görüşlerine göre anlamaya çalışırlar. Ahlâk bağlamı araştırmacının, en üst düzeyde etik davranış standartlarına uymasını ister. Bilimsel bütünlüğün açık şekilde ihlal edildiği başlıca durumlar şunlardır: Verilerin uydurulması, aşırmacılık, araştırma bulgularının seçilerek yayımlanması, araştırmaya önemli katkısı olan kişilerden bahsedilmemesi, araştırma fonlarının kötüye kullanılması ve insan ve hayvan katılımcılara etik dışı davranmak. APA'nın Etik Yönetmeliği, araştırmacılara rehberlik eder ve araştırmalarda aldatma ve hayvan kullanma gibi konulardaki risk/fayda ikilemlerini değerlendirebilmeleri için onlara yardımcı olur.

Araştırmacılar, davranış ve zihinsel süreçlere ilişkin iddialarda kuşkucu olmalıdır. Bilim insanları tüm araştırma bulgularının olasılıklı olduğunu bilirler ama yine de bir iddia için bulunabilecek en güçlü kanıtlar, birçok araştırmadan elde edilen birbirleriyle uyumlu kanıtlardır. Araştırmaların medyada yayımlanması, başlıca üç şekilde soruna yol açar: Medya raporları bilimsel araştırmaya dayanmayabilir; veya yayımlanan araştırmanın standardı düşük olabilir veyahut tipik olarak bu yayınlar orijinal araştırmanın özetleridir. Medyadaki bilimsel yayınları değerlendirmenin önemli bir ilk adımı, araştırmanın yöntemi ve işlemleri hakkında daha fazla bilgi edinmek için orijinal araştırmaya ulaşmaktır.

Araştırmaya başlamanın ilk adımı, bir araştırma sorusu belirlemektir. Öğrencilerin araştırma sorularına ulaşmalarının başlıca kaynakları ders kitapları, dersler ve öğretim üyeleriyle etkileşimleridir. Bir sonraki adım, bir araştırma hipotezi kurmaktır. Bir araştırma hipotezi, test edilecek olay hakkında yapılan geçici bir açıklamadır ve çoğunlukla, öngörülen sonuç hakkında yapılan bir açıklamayı da içeren bir kestirimdir. Araştırma hipotezleri birçok şekilde kurulabilir ama bunun temel bir adımı, araştırılmak istenilen konuyla ilişkili literatürün incelenmesidir. Nihayet, bir araştırma sorusuna verilen cevapların, psikologların davranışı ve zihinsel süreçleri anlamalarına anlamlı bir katkısının olup olmadığını değerlendirmek de önemlidir.

Çok yöntemli yaklaşım, araştırma sorusunu cevaplandırmak ve davranış hakkında daha yetkin bir anlayış kazanmak için çeşitli araştırma yöntemleri ve ölçümleri kullanır. Bilim insanları bilirler ki davranışın bir tek yöntemle ölçümü kusurlu ve eksik bir ölçümdür; çok yöntemli ölçüm, bir yöntemin yarattığı bir boşluğu, diğer yöntem(ler)in telafi etmesine olanak sağlar. Elinizdeki kitabın amacı psikologların, sorularını cevaplamak için kullandıkları çeşitli araştırma yöntemlerini size tanıtmaktır.

TEMEL KAVRAMLAR

bilimsel yöntem 4
görgül yaklaşım 7
etnik merkezcilik yanlılığı 11
hipotez 21
çok yöntemli yaklaşım 23

GÖZDEN GEÇİRME SORULARI

1 Bilimsel yöntemin iki önemli özelliğini betimleyin.
2 Önceki psikologlar psikoloji araştırmaları için neden en çok görgül yaklaşımı seçtiler?
3 Bilgisayar, psikolojinin 20'nci yüzyıldaki gelişiminde birçok şekilde etkili olmuştur; bunlardan ikisini söyleyin.
4 a) Sosyal ve kültürel faktörlerin, psikologun araştırma konusunu seçmesini nasıl etkileyebileceğine ve b) sosyal ve kültürel faktörlerin, toplumun araştırma bulgularını kabul etmesini nasıl etkileyebileceğine birer örnek verin.
5 Etnik merkezci yanlılığın bir araştırmada nasıl sorun yaratabileceğini ve araştırmacıların bu yanlılıktan nasıl kaçınabileceklerine dair bir öneride bulunun.
6 Araştırmanın, bir "ahlâkî bağlam"da yapılması ne demektir?
7 Araştırma yaparken araştırmacıların karşılaşabilecekleri iki etik ikilem açıklayın.
8 Araştırmacıların araştırma bulguları hakkında kuşkucu olmalarının nedenini ve onların tutumlarının, alan dışındaki diğer insanların tutumundan nasıl farklı olabileceğini açıklayın.
9 Medyada yayımlanan araştırma bulgularını (örneğin, kendini geliştirme kitapları, televizyon, dergi) okuyan başka birisinin, neden bu bulguları eleştirel bir gözle değerlendirmesi gerektiğine dair üç neden söyleyin.
10 Bir araştırmaya başlarken, araştırmacıların attıkları ilk üç adım hangisidir?
11 Bir araştırmaya başlarken psikoloji literatürünü incelemenin önemini gösteren iki neden söyleyin.
12 Araştırmada çok yöntemli yaklaşım kullanmayı açıklayın ve bu yaklaşımın en önemli avantajını söyleyin.

DÜŞÜNDÜRÜCÜ SORULAR

1 Videoda şiddet oyunu oynamanın, televizyonda pasif şekilde şiddet filmi izlemeye kıyasla, insanların daha saldırgan olmalarına neden olduğu hipotezini ele alın.
   - A Bu hipotezi nasıl test edebilirsiniz? Yani, farklı şekillerde şiddete maruz kalmanın yol açtığı iki farklı şiddet deneyimini karşılaştırmak için ne yapabilirsiniz?
   - B Şiddete maruz kaldıktan sonra insanların saldırgan bir şekilde davranıp davranmadıklarını nasıl belirlersiniz?
   - C Başka faktörlerin değil de şiddete maruz kalmanın saldırgan davranışa neden olduğundan emin olmak için, başka hangi faktörleri göz önüne almanız gerekir?

2 Araştırmacılar, psikoloji kavramları hakkında çıkarımlarda bulunmak için davranışa ilişkin kendi gözlemlerini kullanırlar. Örneğin, "can sıkıntısı", bir kimsenin oturduğu sandalyede kaç defa yer değiştirdiği (kıpırdadığı) sayılarak ölçülebilir.
   - A Psikologların kullandıkları şu kavramları değerlendirmek için gözleyebileceğiniz bir davranış bulun: *Kişiler arası çekicilik*, *korku*, *hoşlanma* ve *utangaçlık*.
   - B A için bulduğunuz davranışların her birisi, başka hangi ölçülebilir kavramın göstergesi olabilir? Örneğin, sandalyede kıpırdanmak can sıkıntısından ziyade heyecanın veya endişenin bir ölçütü olabilir.

3 Etnik merkezci yanlılığın şu araştırma sorularını nasıl etkileyebileceğini belirleyin, sonra da bu yanlılığı azaltmak için öneride bulunun.
   - A Bir araştırmacı mutluluğun, tatmin duygusuyla (yani, kişisel hedeflerini en üst düzeye çıkarma yeteneğiyle) ilişkili olup olmadığını belirlemeye çalışıyor ve Çin'de ve ABD'de mutluluk ile kendini gerçekleştirme arasındaki ilişkiyi karşılaştırıyor.
   - B Bir psikolog, yaşlı yetişkinlerdeki romantik ilişkileri araştırıyor. Araştırmacı partnerlerin buluştukları zaman nasıl etkileşime girdiklerine odaklanmış olmakla birlikte, birbirleriyle nasıl mesajlaştıklarını da araştırmak istiyor. Özel olarak, partnerlerin birbirleriyle mesajlaşmalarının, daha yüzeysel romantik ilişkilere neden olacağı düşüncesini test etmek istiyor.
   - C Bir araştırmacı, insanların terör tehdidine karşı tepkileri ile ve özellikle de terörizmin insanların kendi ölümlülük duygularını harekete geçirdiği teorisiyle ilgileniyor. Araştırma önerisinde, insanlara öteki dünyayı hatırlatmanın, onların terör korkusunu azaltacağını öne sürüyor.

4 Aşağıdaki sütunların ilkinden seçeceğiniz bir olay veya davranışı, ikincisinden seçeceğiniz bir sonuca bağlayarak bir hipotez kurun, ve daha sonra üçüncü sütundan mümkün bir açıklama belirleyin. Her bir olay, sonuç ve açıklamayı sadece bir defa kullanın. Değişkenlerin birden fazla kombinasyonu da mümkün olabilir.

Olay veya Davranış	Sonuç	Açıklama
1 yoksulluk içinde yaşama	1 yardımda düşüş	1 fizyolojik uyarılma, öfke
2 ırkçı ayrım	2 risk alıcı davranış	2 belirsizlik deneyimi
3 olumsuz kalıp yargısı olanlarla karşılaşma	3 psikolojik sıkıntı	3 başkalarının yardımseverliğini algılamada düşüş
4 videoda şiddet oyunu oynama	4 başkasıyla etkileşim isteğinde artış	4 endişede azalma

İleri Alıştırmaların Cevapları

1 Zayıf beden fotoğrafları görmek, bedensel tatminsizliğe neden olur; muhtemelen bunun nedeni, kendi bedenini ideal bir bedenle karşılaştırmaktır [Dittmar, Halliwell ve Ive, 2006).
2 On bir Eylül 2001 terör saldırısından sonra ölümlü trafik olayları arttı; muhtemelen bunun nedeni, hava yoluyla ulaşımdan korkmaktır [Gigerenzer, 2004].
3 İnsanların davranışlarını ve duruşlarını taklit etme, onlara daha fazla yardım yapılmasına neden olur; muhtemelen, empati duygusundaki artış nedeniyle [van Baaren, Holland, Kawakami ve van Knippenberg, 2004].

Düşündürücü Soruların Cevapları

A Bu hipotezi test etmenin bir yolu, iki grup katılımcı oluşturup birisine videoda şiddet oyunu oynatmak, diğerine ise televizyonda şiddet filmi izletmektir. İkinci bir test, her iki gruba her iki tip şiddeti farklı zamanlarda uygulamaktır.

B Insanların video oyunundan sonra mı yoksa televizyondan sonra mı daha saldırgan olduklarını belirlemek için, bazı saldırgan davranış ölçütlerinizin olması gerekir. Potansiyel olarak, sınırsız sayıda böyle ölçüm var; eğer araştırmacının yeteneği bu sayıyı sınırlamazsa tabi. İyi bir ilk adım, başka araştırmacıların kullanmış oldukları ölçütleri kullanmaktır; böylece, elde ettiğiniz sonuçları önceki araştırma sonuçlarıyla karşılaştırabilirsiniz. Saldırganlık ölçütleri arasında örneğin, insanlara öfke yaratacak hayalî durumlarda nasıl davranacaklarını işaretlemeleri veya şiddete maruz kaldıktan sonra deneyciye (veya başkalarına) nasıl davrandıklarını gözlemektir. İkinci durumda, katılımcıların şiddet içeren (veya içermeyen) davranışlarını kaydetmek için bir kontrol listesinin veya başka yöntemlerin olması gerekir.Unutmayınız ki saldırganlık çok çeşitli şekillerde tanımlanabilir; örneğin, fiziksel davranış, sözel davranış ve hatta düşünce şeklinde olabilir (ama bu son durumda şiddeti ölçmenin zorluğunu da gözden kaçırmayın).

C İki grubun, televizyon ve video farkı *hariç* her yönden benzer olduklarından emin olmak önemlidir. Örneğin, araştırmanızda iki grup katılımcı olduğunu varsayın: Bir grup televizyon izledi, diğeri ise video oyunu oynadı. Yine varsayın ki sizin sonuçlarınız, sizin saldırganlık ölçeğiniz üzerinde, video oynayan grubun televizyon seyreden gruptan daha saldırgan olduğunu gösterdi.

Eğer video grubunuz başlangıçta zaten televizyon grubundan daha saldırgan olsaydı, o zaman bir problem ortaya çıkardı. Araştırmanızda gözlediğiniz saldırganlık farkının nedeninin, video ve televizyonun farkı mı, yoksa gruplarınızın doğal saldırganlık farkı mı olduğunu bilemezsiniz. Bu nedenledir ki her iki gruptaki katılımcılarınızın, şiddete maruz kalmadan önce benzer olduklarından emin olmak istersiniz. Bu kitapta daha sonra, gruplarınızı nasıl denkleştireceğinizi öğreneceksiniz.

Aynı zamanda, katılımcılarınızın deneyimlerinin başka yönlerden benzer olduklarından da emin olmak istersiniz. Örneğin, her iki grubun şiddete maruz kalma sürelerinin benzer olduğundan da emin olmak istersiniz. Ayrıca, televizyon programındaki şiddetin derecesinin, videodaki şiddetin derecesine benzer olmasını da sağlamaya çalışırsınız. Katılımcıların deneyimlerinin, başka faktörler açısından farklı olmamaları da önemlidir; örneğin, şiddete maruz kalma esnasında başka insanların olup olmaması ve günün hangi saatinde şiddete maruz kalındığı gibi. Video oyununun televizyon izlemekten daha fazla (veya daha az) saldırganlığa neden olduğunu göstermek için en önemli nokta şudur ki gruplar arasındaki farkın tek nedeni, şiddete maruz kalma şekli olmalıdır.

BÖLÜM İKİ

Bilimsel Yöntem

BÖLÜMÜN ANA HATLARI

BİLGİYE BİLİMSEL YAKLAŞIM VE GÜNLÜK YAKLAŞIM
Genel Yaklaşım ve Tutum
Gözlem
Kavramlar
Rapor Etme
Araçlar
Ölçme
Hipotezler
BİLİMSEL YÖNTEMİN AMAÇLARI
Betimleme
Kestirim
Açıklama
Uygulama
BİLİMSEL TEORİLERİN KURULMASI VE TEST EDİLMESİ
ÖZET

BİLGİYE BİLİMSEL YAKLAŞIM VE GÜNLÜK YAKLAŞIM

- Bilimsel yöntem görgüldür (empirical) ve sistemli, kontrollü gözlem gerektirir.
- Bilim insanı bir deney yaptığı zaman, en üst düzeyde kontrol olanağı elde eder; bir deneyde araştırmacı bağımsız değişkenleri değişimler ve bunların davranışa olan etkisini belirlemeyi amaçlar.
- Bağımlı değişkenler davranışın ölçüsüdür ve bağımsız değişkenlerin etkisini değerlendirmek için kullanılır.
- Bilimsel raporlar yansızdır ve nesneldir; işevuruk (operasyonel) tanımlar kullanıldığı zaman kavramlar açık şekilde anlaşılabilir ve iletilebilir.
- Bilimsel araçlar doğrudur ve hassastır; fiziksel ve psikolojik ölçekler geçerli ve güvenilir olmalıdır.
- Bir hipotez, bir olayın geçici bir açıklamasıdır; test edilebilir hipotezler açık şekilde tanımlanmış (işevuruk) kavramlar kullanır ve kısır döngü içermez, ve gözlenebilir kavramlar kullanır.

Bilimsel yöntem, yüz yıldan fazla bir zamandır psikoloji alanında kullanılmaktadır. Bilimsel yöntem özel tipte bir donanım gerektirmediği gibi, belirli bir işlemle veya teknikle de bağlantılı değildir. Bölüm 1'de betimlendiği gibi bilimsel yöntem, bilim insanlarının soru sormalarını ve bu sorulara cevap bulmak için kullandıkları yöntemleri ve akıl yürütme şekillerini kapsar. Kendimiz ve içinde yaşadığımız dünya hakkında bilgi edinmenin birçok verimli şekli vardır; örneğin, felsefe, din bilim, edebiyat, sanat ve diğer disiplinler. Bilgi edinmenin bir şekli olarak bilimsel yöntemi anlamanın en iyi yollarından birisi, onu "günlük hayatta" bilgi edinme yöntemlerimizden ayırt etmektir. Nasıl ki teleskop ve mikroskop günlük hayattaki görme yeteneklerimizi genişletiyorsa, bilimsel yöntem de günlük hayattaki bilgi edinme olanaklarımızı genişletmektedir.

Bilgi edinmenin bu iki şekli arasındaki başlıca farklar Tablo 2.1'de özetlenmiştir ve sonraki sayfalarda kısaca açıklanmıştır. "Bilimsel" başlığı altında listelenmiş olan özelliklerin hepsi birlikte, bilimsel yöntemi tanımlamaktadır.

TABLO 2.1 BİLGİYE BİLİMSEL VE BİLİMSEL OLMAYAN (GÜNLÜK) YAKLAŞIMLARIN ÖZELLİKLERİ*

	Bilimsel olmayan (günlük))	Bilimsel
Genel yaklaşım:	Sezgisel	Görgül
Tutum:	Eleştirel değil, kabullenme	Eleştirel, kuşkucu
Gözlem:	Rastgele, kontrolsüz	Sistemli, kontrollü
Kavramlar:	Belirsiz, çok anlamlı	Açık tanımlı, işevuruk, belirli
Bildirme:	Yanlı, öznel	Yansız, nesnel
Araçlar:	Yanlış, hassas değil	Doğru, hassas
Ölçekler:	Geçerli veya güvenilir değil	Geçerli ve güvenilir
Hipotezler:	Test edilemez	Test edilebilir

*Kısmen, Marx'ın (1963) yaptığı ayrıma dayanmaktadır.

Genel Yaklaşım ve Tutum

Bölüm 1'de, bir araştırmacı gibi düşünmek için kuşkucu olmanız gerektiğini söyledik. Psikoloji alanında çalışan bilim insanları davranış ve zihinsel süreçlere ilişkin iddiaları kabul etmede çok temkinlidir ve bunları kabul etmeden önce kanıtları eleştirel bir gözle değerlendirirler. Böyle olmakla birlikte, günlük düşünme şeklimizde iddialara ait kanıtları ya hiç değerlendirmeyiz ya da çok az değerlendiririz. Genellikle, günlük hayattaki kararlarımızın çoğunu sezgilerimize dayanarak veririz. Çoğunlukla bunun anlamı şudur ki eylemlerimizi yöneten şey "doğru olduğu kanaatinde olmamız" veya "akla uygun görünmesi"dir. Sezgi, başka bir bilgimizin olmadığı durumlarda değerli olabilir ama her zaman doğru değildir.

Karar vermek için sezgilerimize dayandığımız zaman, algılarımızın bilişsel yanlılıklarla çarpıtılmış olabileceğini veya bütün kanıtları göz önüne almamış olabileceğimizi çoğunlukla fark etmeyiz (bk., Kutu 2.1; Kahneman ve Tversky, 1973; Tversky ve Kahneman, 1974). Psikologlar bilimsel yöntemi kullanarak, *doğrulama yanlılığı*ndan (confirmation bias) kaçınmaya çalışırlar –doğrulama yanlılığı doğal bir eğilimimizdir: Bizim sezgilerimizle, beklentilerimizle ve mevcut anlayışımızla uyumlu olan kanıtları bulmaya çalışma, ama bu esnada, beklentilerimize, sezgilerimize uymayan kanıtları ise görmezden gelme veya inkâr etme eğilimidir. Doğrulama yanlılığı, insanların haber programı (örneğin, Fox, CNN, MSNBC gibi) seçimlerini yönetir ve onları önceki tutumları, inançları ve davranışları ile uyuşmayan bilgilerden kaçınmaya teşvik eder; bu kaçınılan bilgi onları yanlış sonuçlara götürse bile böyle olabilir (Hart ve ark., 2009).

Psikoloji araştırmaları, doğrulama yanlılığının örnekleriyle doludur. Örneğin, küçük çocuklar bile *aldatıcı korelasyon* (illusory correlation) denilen bilişsel hataya duyarlıdır; bu korelasyon, olaylar arasında olmayan ilişkileri algılama eğilimidir. Bir araştırmada (Susskind, 2003) araştırmacılar çocuklara, birçok kadın ve erkek fotoğrafı gösterdiler; fotoğraflardaki insanlar ya alışıldık (bir kadının örgü örmesi gibi), ya alışılmadık (bir erkeğin örgü örmesi gibi), ya da nötr (kitap okuma gibi) işler yapıyorlardı. Çocukların hangi tip fotoğrafı kaç defa gördüklerini tahmin etmeleri istenildiği zaman, alışıldık davranışları olduğundan daha fazla sayıda

KUTU 2.1

SANDIĞIMIZ KADAR ZEKİ (VEYA İYİ) DEĞİLİZ

Psikologlar uzun zaman önce fark ettiler ki insan düşüncesi hataya ve yanlı olmaya eğilimlidir. Bu yanlılıklar bizim nasıl kestirimde bulunacağımızı, nasıl karar vereceğimizi ve başkalarıyla nasıl etkileşimde bulunacağımızı etkiler. Genel okuyucuyu hedef alan bazı kitaplar bu araştırmaları konu almaktadır ve bu kitapların bazıları, en çok satan kitaplar listesine girmiştir. Şu kitapları inceleyebilirsiniz:

- *Blind Spot: Hidden Biases of Good People* by Mahzarin R. Banaji and Anthony G. Greenwald (Delacorte, 2013)
- *The Invisible Gorilla: How Our Intuitions Deceive Us* by Christopher Chabris and Daniel Simons (Crown, 2010)
- *The Signal and the Noise: Why So Many Predictions Fail—But Some Don't* by Nate Silver (The Penguin Press, 2012)
- *Thinking, Fast and Slow* by Daniel Kahneman (Farrar, Straus, and Giroux, 2011)
- *You Are Not So Smart: Why Your Memory Is Mostly Fiction, Why You Have Too Many Friends on Facebook, and 46 Other Ways You're Deluding Yourself* by David McRaney (Gotham, 2011)

gördüklerini bildirdiler. Bu çocuklar kendi beklentileriyle uyumlu olan fotoğraflara, beklemediklerinden daha fazla dikkat ederek, kadınlara ve erkelere dair yerleşik anlayışlarını doğrulamış oldular.

Bilgiye bilimsel yaklaşım, sezgisel olmaktan ziyade görgüldür (empirical). Görgül bir yaklaşım, sorulara cevap vermek için *doğrudan gözleme* ve *deney yapma*ya önem verir. Bu, sezginin bilimde bir rolünün olmadığı anlamına gelmez. Başlangıçta araştırmayı, bilim insanının sezgileri yönlendirir. Böyle olmakla birlikte eninde sonunda bilim insanı, doğrudan gözlemin ve deneyin sağladığı görgül kanıtlara bağlı olmak ister.

Gözlem

Sadece başkalarının yaptıklarını gözleyerek, davranış hakkında pek çok şey öğrenebiliriz. Ama günlük gözlemler daima dikkatli ve sistemli bir şekilde yapılmaz. Birçok insan, gözlediği olayları etkileyebilecek faktörleri kontrol etmeye veya bertaraf etmeye çalışmaz. Böyle olunca da rastgele günlük gözlemlerimize dayanarak, çoğunlukla yanlış sonuçlara varırız. Örneğin, artık klasik olan Akıllı Hans olayını ele alalım. Hans bir attır ve bir matematik öğretmeni olan sahibi, onun inanılmaz bir yeteneğinin olduğunu söylemektedir. Hans sayabilmekte, basit toplama ve çıkarma işlemlerini (kesirli sayılar dâhil) yapabilmekte, Almanca okuyabilmekte, basit soruları ("Bu hanımın elinde ne var?" gibi) cevaplayabilmekte, tarih ve zaman söyleyebilmektedir (Watson, 1914/1967). Hans soruları cevaplamak için ön ayaklarını yere vuruyor veya kendisine gösterilen çeşitli seçenekler arasından birisini burnuyla işaret ediyordu. Sahibi, bu atın gerçekten zeki olduğunu düşünüyordu ve onun davranışlarını etkileyen bir hile yaptığını reddediyordu. Ve gerçekten de Hans, soruları sadece sahibi değil, başkaları sorduğu zaman da cevaplayabiliyordu.

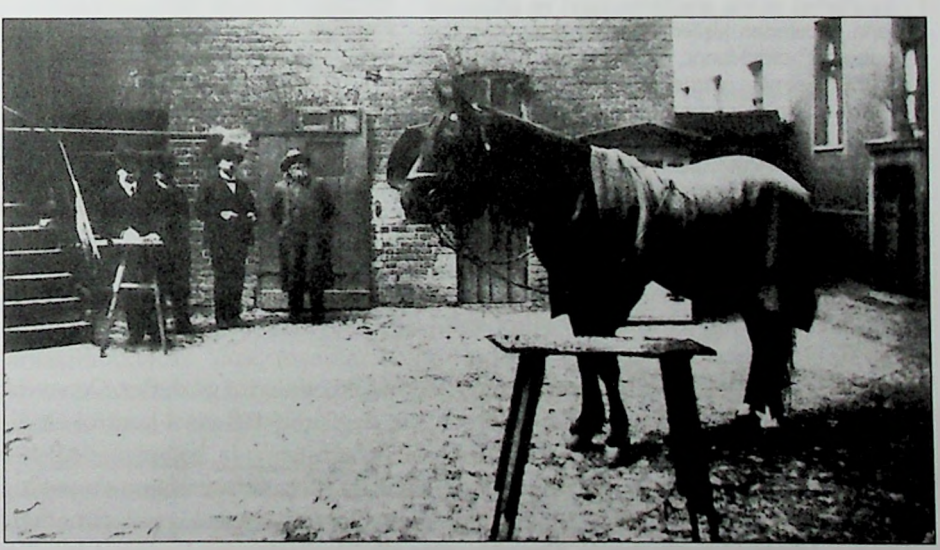
Gazeteler Hans'ın yaptıklarına dair açıklamalar yazdılar, ve yüzlerce insan, bu inanılmaz atı görmeye geldi (Şekil 2.1). Hansın yeteneklerinin temellerini keşfetmesi için, 1904 yılında bir bilim komisyonu oluşturuldu. Sahibi adına ne yazık ki bu komisyon, Hans'ın iki durumda zeki bir davranış gösteremediğini buldu. İlk olarak, soruyu soran kişi sorunun cevabını bilmediği durumda, Hans da cevabı bilemedi. İkinci olarak, soruyu soran kişiyi tanımadığı durumda da Hans soruyu cevaplayamadı. Komisyondaki bilim insanları ne gözlediler? Hans'ın soruya değil, soruyu soran kişinin küçük hareketlerine tepki verdiğini fark ettiler. Soruyu soranın hafifçe öne eğilmesi Hans'ın yere tekme vurma hareketini başlatıyordu ve yukarıya veya geriye doğru yapılan herhangi bir hareket ise, yere vurma hareketini sona erdiriyordu. Komisyon gösterdi ki Hans ön ayağını yere vururken veya başıyla işaret ederken, soruyu soranlar farkında olmadan kendisine ipucu veriyorlar. Bu çerçevede diyebiliriz ki Hans, kendisini gözleyen birçok insandan daha iyi bir gözlemciydi.

Temel Kavram

Akıllı Hans'ın bu ünlü açıklaması gösteriyor ki bilimsel gözlem (günlük gözlemden farklı olarak) sistematiktir ve kontrollüdür. Nitekim **kontrol**ün, bilimin en temel ögesi olduğu, bilimi, bilimsel olmayan işlemlerden ayırt ettiği öne sürülür (Boring, 1954; Marx, 1963). Akıllı Hans olayında araştırmacılar, her defasında bir koşulu değişimleyerek, kontrol sağladılar; örneğin, soruyu soran kişinin cevabı bildiği ve bilmediği durumlarda ve Hans'ın soruyu soranı görüp görmediği durumlarda onun tepkilerini gözlediler (bk., Şekil 2.1). Bilim insanları kontrollü gözlem kullanarak, bir olayı meydana getiren faktörleri açık bir şekilde ortaya

ŞEKİL 2.1 Üstte: Akıllı Hans, izleyicilerin önünde gösteri yapıyor. Altta: Hans kontrollü koşullar altında, soruyu soran kişiyi görmediği durumda, test ediliyor





çıkardılar. Akıllı Hans'ın dikkatli ve sistematik gözlemi, bilim insanlarının davranış hakkında anlayış kazanmak için kullandıkları kontrolün bir örneğidir. Kutu 2.2'de, 100 yıldan fazla bir zaman önce yaşanmış olan akıllı Hans'ın hikâyesinin, bugün bile bilim insanlarına nasıl bilgi verdiğinin bir örneğini okuyacaksınız.

Temel Kavram

Bilim insanları en üst düzeydeki bir kontrolü, bir deney yaptıkları zaman elde ederler. Bir **deney**de bilim insanları, bir veya daha fazla sayıda faktörü değişimlerler

KUTU 2.2

KÖPEKLER KANSERİ TEŞHİS EDEBİLİR Mİ?

Kanser belirleme yöntemleri üzerine yapılan araştırmalar, *British Medical Journal*'da 2004 yılında yayımlanan bir araştırma ile ilginç bir dönüşüme uğradı. Bu araştırmada köpeklerin, bir kimsenin idrarını koklayarak mesane kanserini güvenilir bir düzeyde başarıyla belirledikleri gösterildi (Willis ve ark., 2004). Bu araştırmadan sonra çok sayıda hikâye ortalıkta dolaştı. Örneğin, köpek sahipleri, kendilerine kanser teşhisi konulmasından önce köpeklerinin aniden aşırı koruyucu veya derideki yaralarla uğraşır hâle geldiklerini bildirdiler. Bu gibi hikâyelere olan ilgi öyle büyüktü ki *60 Dakika* gibi televizyon programlarında buna benzer gösteriler yapıldı. Ama kuşkucular Akıllı Hans olayını anımsattılar ve köpeğin, araştırmacının verdiği küçük ipuçlarını kullanarak kanserli ve kontrol örneklemlerini (idrarlarını) ayırt ettiğini öne sürdüler. Araştırmanın taraftarları ısrarla araştırmacıların ve gözlemcilerin, incelenen kişilerin gerçek durumlarından habersiz olduklarını, bu nedenle köpeklere ipucu vermiş olamayacaklarını iddia ettiler. Daha sonraki araştırmalar ise karışık sonuçlar buldu (örneğin, Gordon ve ark., 2008; McCulloch ve ark., 2006 gibi). Bu yeni kanser teşhisi alanındaki araştırmacılar, daha fazla araştırma yapabilmek için araştırma fonlarına başvurdular. Şimdi biz, köpeklerin gerçekten kanseri teşhis edip edemediklerini öğrenmek için, bu zorlu araştırmaların sonuçlarını bekliyoruz.



Temel Kavram

ve bu değişimlemenin davranış üzerindeki etkilerini gözlerler. Deneycinin, davranış üzerindeki etkisini belirlemek için değişimlediği veya kontrol ettiği faktörlere **bağımsız değişken**ler[1] denir. En basit bir araştırmada, bağımsız değişkenin iki düzeyi vardır. Bu iki düzey, çoğunlukla, bazı işlemlerin varlığını ve yokluğunu temsil eder. İşlemin olduğu koşula yaygın şekilde deneysel koşul, olmadığı koşula ise kontrol koşulu denilir. Örneğin, eğer alkol kullanmanın karmaşık bilgilerin hızlı ve doğru işlenmesi yeteneğine olan etkisini araştırmak istiyorsak, bağımsız değişken,

[1]Bağımsız değişkenin düzeyleri bazen bir araştırmacı tarafından seçilir. Bir *bireysel farklar değişkeni*, bireyler arasında değişen bir özellik veya niteliktir. Örneğin, katılımcıların cinsiyeti (kadın, erkek) bir bireysel farklar değişkenidir. Araştırmacılar davranışın cinsiyete göre değişip değişmediğini araştırdıkları zaman kadın ve erkek katılımcılar seçerler ve bu faktörü bir bireysel farklar değişkeni olarak incelerler. Bölüm 6'da göreceğimiz gibi, değişimlenen ve seçilen bağımsız değişkenler arasında önemli farklar vardır.

bir içeceğin içinde alkolün olması veya olmaması olur. Deneysel koşuldaki katılımcılar bir içecek içinde alkol alır, kontrol koşulundaki katılımcılar ise aynı içeceği alkolsüz içer. Bu bağımsız değişkeni değişimledikten sonra araştırmacı, karmaşık bilgiyi işleyip işleyemediklerini görmek için, katılımcıların karmaşık bir video oyununu oynamalarını isteyebilir.

Temel Kavram

Bağımsız değişkenin etkisini (eğer varsa tabi) değerlendirmek için kullanılan davranış ölçümlerine, **bağımlı değişken**ler denir. Bizim verdiğimiz alkolün karmaşık bilginin işlenmesine olan etkisi araştırmasında araştırmacı, deney ve kontrol gruplarının zor bir video oyunu oynarken yaptıkları hataların sayısını belirleyebilir. Böylece bu araştırmadaki hata sayısı, bağımlı değişken olur.

Bilim insanları davranıştaki herhangi bir farka (bağımlı değişken), bağımsız değişkenin farklı koşullarının neden olup olmadığını belirlemeye çalışırlar. Bizim örneğimizde bu demektir ki video oyunu oynarken iki grup arasında gözlenen bir hata farkının nedeni, bağımsız değişkenin farklı koşullarıdır –alkolün olduğu ve olmadığı koşullar. Bununla birlikte, bu açık sonuca varabilmek için bilim insanlarının kontrol tekniklerini doğru şekilde uygulamış olmaları gerekir. Bu kitabın her bir bölümü, araştırmacıların davranışı ve zihni araştırmak için kontrol tekniklerini nasıl kullandıklarını vurgulayacaktır.

Kavramlar

Kavram terimini, canlı ve cansız şeyleri, olayları ve şeyler ve olaylar arasındaki ilişkileri ve bunların özelliklerini kast etmek için kullanıyoruz (Marx, 1963). "Köpek" bir kavramdır; "havlama" ve "itaat" de birer kavramdır. Kavramlar semboldür ve biz kavramlarla iletişimde bulunuruz. Düşüncelerin açık, tam iletişimi, kavramlarımızı açık şekilde tanımlamamızı gerektirir.

Günlük konuşmada çoğunlukla, bir kavramı nasıl tanımladığımızı fazla mesele yapmayız. Örneğin, her iki taraf da kelimelerin anlamını tam olarak bilmese de, birçok kelime yaygın şekilde kullanılmakta ve görünüşe bakılırsa, anlaşılmaktadır. Yani, insanlar çoğunlukla bir başkasıyla, konuştukları şeyin tam olarak farkında olmadan iletişimde bulunurlar. Bu, biraz gülünç gelebilir ama ne demek istediğimizi anlamak istiyorsanız, şunu yapmayı deneyebilirsiniz: Birkaç kişiye, zekâda kalıtımın mı yoksa çevrenin mi payının daha fazla olduğuna inandıklarını sorun. Zekânın temellerini tartıştıktan sonra kendilerine, "zekâ" ile ne kast ettiklerini sorun. Büyük olasılıkla, zekânın kaynağını tartıştıktan sonra bile insanların zekâyı tanımlamada zorlandığını ve farklı tanımlar yaptıklarını göreceksiniz. Açıktır ki zekânın kaynağının kalıtım mı yoksa çevre mi olduğu sorusunu cevaplayabilmek için, tüm tarafların kabul edebilecekleri kesin bir tanıma ihtiyacımız var.

Temel Kavram

Psikolojik bilimlerde "kavram"ın araştırılması o kadar önemlidir ki araştırmacılar kavram için farklı bir isim kullanırlar: Yapılar. Bir yapı, bir kavram veya bir fikirdir; zekâ, depresyon, saldırganlık ve bellek psikolojik yapılardan bazılarıdır. Bir bilim insanının bir yapıya anlam vermesinin bir şekli, işevuruk

İLERİ ALIŞTIRMA

Bu alıştırmada, önce bir araştırmanın aşağıdaki özetini okumanızı, sonra özetin sonundaki iki soruyu cevaplamanızı istiyoruz.

Psikolojinin nispeten yeni bir alanı olan "pozitif psikoloji" olumlu duyguları, olumlu karakter özelliklerini, ve olumlu kurumları araştırmaya odaklanır. Pozitif psikolojide araştırmanın amacı, insanların kendilerini iyi hissetmelerini ve mutlu olmalarını destekleyecek davranış biçimlerini belirlemektir (Seligman, Steen, Park ve Peterson, 2005). Bu alandaki çalışmaların bazıları da *minnet* duygusunu araştırmaktadır; kendisine değerli bir şey veren bir insana karşı duyulan olumlu duygu (Bartlett ve DeSteno, 2006). Bazı araştırmalar, minnet duyan insanların olumlu sosyal davranış gösterme olasılıklarının daha yüksek olduğunu düşündürmektedir –yani, başkalarına faydalı olacak şekilde davranmak.

Bartlett ve DeStefano (2006) *işbirlikçi* (deneycinin gizlice anlaştığı ve onun istediği gibi davranan sözde katılımcı; bk., Bölüm, 4) kullanarak, minnet duygusu ile başkalarına yardım davranışı arasındaki ilişkiyi araştırdılar. Her bir katılımcı ilkin, el-göz koordinasyonu gerektiren uzun ve sıkıcı bir görevi yapmak için bir işbirlikçi ile ekip oluşturdu. Daha sonra, katılımcıların üçte birinin bilgisayar ekranı silindi ve kendilerinden, yapmış oldukları işi tekrar yapmaları istendi. Ama işbirlikçi ekranı düzelterek problemi çözdü ve katılımcıyı işi tekrar yapma zahmetinden kurtardı; böylece, katılımcıda bir minnet duygusu uyandırmış oldu. Diğer katılımcılar için durum farklıydı. Görev tamamlandıktan sonra, katılımcıların diğer üçte biri işbirlikçi ile birlikte eğlenceli bir video oyunu oynadı (olumlu duygu) ve kalan üçte biri ise işbirlikçi ile kısa bir sohbet yaptı (nötr duygu). Bazı soru listelerini doldurduktan sonra işbirlikçi her bir katılımcıdan, kendisinin bir dersine ait çok uzun bir anketi hatır için cevaplamasını istedi. Araştırmacılar, minnet koşulundaki katılımcıların anketi doldurmak için harcadıkları zamanın (ort: 20.94 dakika) olumlu duygu koşulundaki (ort: 12.11 dakika) ve nötr duygu koşulundaki (ort: 14.49 dakika) katılımcılardan daha uzun olduğunu buldular.

1 Bu araştırmadaki bağımsız değişkeni (düzeyi dâhil) ve bağımlı değişkeni belirleyin.
2 Katılımcıların işbirlikçilere daha çok yardım etmelerini sağlayan duygunun basit bir olumlu duygu değil de *minnet* duygusu olduğunu, araştırmacılar nasıl belirleyebildi?

Temel Kavram

(operational) şekilde tanımlamaktır. **İşevuruk tanımlama** bir kavramı sadece, onu üretmek ve ölçmek için kullanılan gözlenebilir işlemlere göre açıklar. Örneğin zekâ, mantıksal ilişkilerin anlaşılmasını öne çıkaran bir kâğıt-kalem testi, kısa süreli bellek ve kelimelerin anlamına aşinalık kullanılarak işevuruk şekilde tanımlanabilir. Bazıları zekânın bu işevuruk tanımını beğenmeyebilir ama bir defa belirli bir test tanımlandıktan sonra, zekânın ne olduğuna dair, en azından *bu tanıma göre*, bir tartışma olmaz. İşevuruk tanımlar, en azından bu tanımların nasıl ve neden kullanıldığını bilenler arasında iletişimi kolaylaştırır.

Her ne kadar işevuruk tanım kesin bir anlam taşıyor ise de kavramların bu iletişim şekli de eleştirilmektedir. Bir problem şudur ki eğer biz zekânın bir işevuruk tanımını beğenmezsek, bir başka işevuruk tanımını yaparız. Peki, bu tanım sayısı kadar çeşitli zekâlar da olduğunu söyleyebilir miyiz? Bu sorunun cevabını ne yazık ki bilmiyoruz. Farklı bir işlemin veya testin yeni bir zekâ tanımı ortaya çıkarıp çıkarmayacağını belirlemek için, daha fazla kanıt elde etmek zorundayız. Örneğin, bir testte yüksek puan alanlar bir diğer testte de yüksek puan alırlar mı? Eğer alırlarsa, ikinci test birinci testle aynı kavramı ölçüyor olabilir.

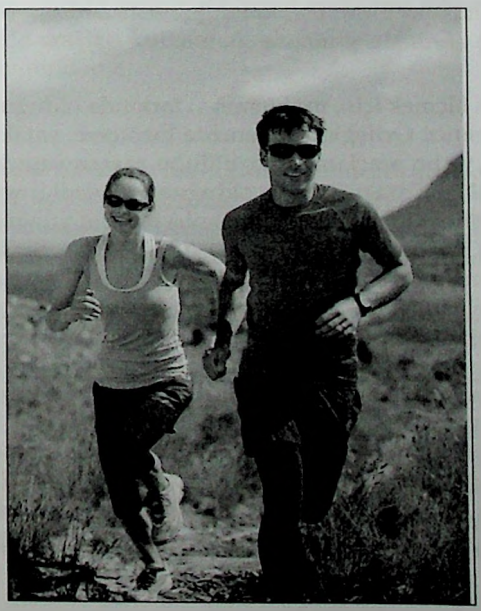
İşevuruk tanım kullanmaya yöneltilen bir diğer eleştiri, tanımların her zaman anlamlı olmadıkları eleştirisidir. Bu eleştiri daha ziyade kültürlerarası araştırmalarla bağlantılıdır; bu araştırmalarda, örneğin, bir kâğıt-kalem zekâ testi, belirli bir kültürel bağlama özgü olan bir bilgiyi harekete geçirebilir. Bir kavramın anlamlı bir şekilde tanımlandığına nasıl karar vereceğiz? Bir defa daha şu noktaya geliyoruz: Çözüm, diğer kanıt türlerine başvurmaktır. Bir testteki performans, zekânın ölçütü olduğu yaygın şekilde kabul edilmiş olan başka görevlerdeki performansla nasıl karşılaştırılabilir? Bilim insanları işevuruk tanımların sınırlılığının genellikle farkındadırlar; ama işevuruk tanımların en güçlü yanı, bilim insanlarının kullandıkları kavramlar üzerinde açık bir şekilde anlaşmalarına yardımcı olmasıdır. İşevuruk tanımın bu gücünün, onun sınırlılıklarından daha üstün olduğu kabul edilmektedir.

Rapor Etme

Birisinden, kaçırdığınız bir ders için size bilgi vermesini istediğinizi varsayın. Muhtemelen, derste yapılanlar hakkında doğru bir bilgi verilmesini istersiniz. Veya, belki de bir partiyi kaçırdınız ve o partide iki arkadaşınız hararetli bir tartışma yaptı ve siz partiye katılan birisinden partide olup bitenleri size anlatmasını istiyorsunuz. Tahmin edebileceğiniz gibi, aldığımız günlük raporlar çoğu zaman kişisel yanlılıklar ve öznel izlenimlerin etkisindedir. Başkalarından bir olayı betimlemelerini istediğiniz zaman, muhtemelen, onların kişisel izlenimlerinin yanı sıra, olayın (her zaman doğru olmayan) detaylarını da öğrenmek istersiniz.

Bilim insanları bulgularını rapor ettikleri zaman, gözlemiş oldukları şeyler ile bu gözlemlerine dayanarak vardıkları sonuçları veya çıkarımları ayırt etmek

ŞEKİL 2.2 Bu sahneyi nasıl betimlersiniz?



isterler. Örneğin, Şekil 2.2'deki fotoğrafı ele alın. Bu fotoğrafta gördüklerinizi bir başkasına nasıl betimlersiniz? Bu fotoğrafı betimlemenin bir şekli, iki kişinin bir yolda koşmakta olduklarını söylemektir. Aynı fotoğrafta, iki kişinin yarış yaptıklarını da söyleyebilirsiniz. Eğer bu ikinci betimlemeyi kullanırsanız, bu demektir ki siz gördüğünüz şeyi değil, gördüğünüz şeyden yaptığınız bir çıkarımı bildiriyorsunuz. Bilimsel bir raporda, iki insanın koşmakta olduğunu rapor etmek tercih edilir.

Rapor etmede betimleme ile çıkarım arasında yapılan bu ayrım, uç bir hâl alabilir. Örneğin, Şekil 2.2'de görülen şeyi koşma olarak betimleme bir çıkarım olarak görülebilir; gerçek gözlemin ise iki insanın ayaklarını yukarı aşağı ve ileriye hızlı ve uzun adımlarla hareketi olduğu söylenebilir. Böyle bir aslına sadık betimleme uygun da olmayabilir. Mesele şudur ki bilimsel raporlarda gözlemcilerin çok çabuk çıkarımda bulunma eğilimine karşı önlem almaları gerekir. Dahası, olaylar yeterince ayrıntılı olmalıdır ama önemsiz detaylara yer verilmemelidir. Gözlem yapmanın ve bu gözlemleri rapor etmenin uygun yöntemleri Bölüm 4'te ele alınacaktır.

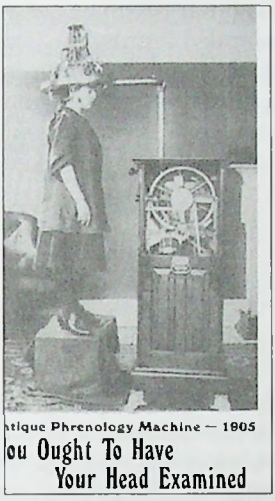
Bilimsel rapor yazma, *yansız* ve *nesnel* olmayı amaçlar. Bir araştırma raporunun yansız olup olmadığını belirlemenin bir yolu, bu raporun bağımsız bir gözlemci tarafından doğrulanıp doğrulanmadığını görmektir. Buna "gözlemciler arası uyum" denir (bk., Bölüm, 4). Ama ne yazık ki birçok yanlılık kolay kolay fark edilemeyecek kadar incelikli olduğu için, bunları bilimsel raporlarda bile keşfetmek zordur. Yumurtaların, yavrular yumurtadan çıkıncaya kadar erkek ebeveynin ağzında kuluçkaya yatırıldığı bir balık türünü ele alın. Bu olayı ilk defa gözleyen bilim insanının, bir süreliğine, erkek ebeveynin yumurtaları yediğini düşünmesi mutlaka anlayışla karşılanmalıdır. Bizim, canlıların ağızlarıyla yapmalarını beklediğimiz şey budur. Ama dikkatli bir gözlemci bekler, beklenmedik sonuçlara hazırlıklı olur ve hiçbir şeye kesin gözüyle bakmaz.

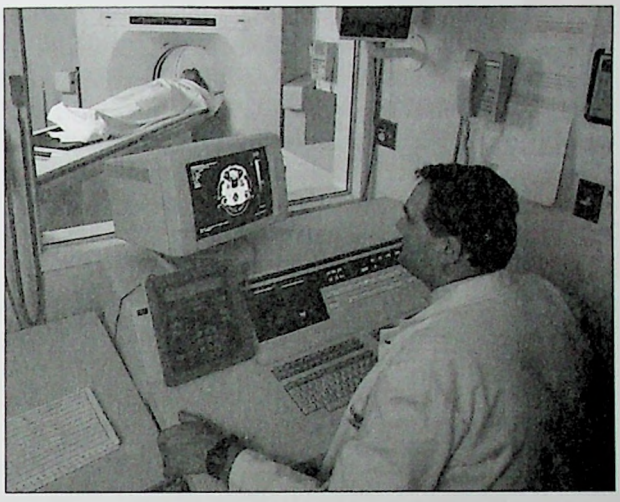
Araçlar

Olayları ölçmek için, muhtemelen farkında olduğunuzdan daha fazla, araçlara dayanırsınız. Örneğin, arabanızda hızölçere, yatak odanızda çalar saate güvenirsiniz ve bu araçlarınız bozulduğu zaman yaşadığınız sıkıntıyı iyi bilirsiniz. *Doğruluk*, bir aracın doğru olduğunu söylediği şey ile doğru olduğu bilinen şey arasındaki fark demektir. Yanlış bir saat yüzünden geç kalırız ve yanlış bir hızölçer yüzünden trafik cezası ödeyebiliriz. Bir aracın doğruluğu onu *ayarlayarak* veya doğru olduğu bilinen bir başka araçla karşılaştırılarak belirlenir. Ayrıca ölçümler, farklı düzeylerde *duyarlılık*ta (precision) yapılabilir. Bir saniyenin yüzde biri düzeyindeki bir ölçüm, onda biri düzeyindeki bir ölçümden daha duyarlıdır.

Davranışı ölçmek için de ölçümlere ihtiyacımız var. Psikolojide kullanılan araçların duyarlılıkta ve doğrulukta, ilk psikoloji laboratuvarının kurulduğu 1879 yılından beri geliştiğini bilmelisiniz. Günümüzde psikoloji alanında, birçok incelikli ölçme aracı kullanılmaktadır (Şekil, 2.3). Psikofizyoloji deneyleri (örneğin, bir kimsenin uyarılma düzeyi değerlendirileceği zaman), kalp atışı ve kan basıncı gibi içsel durumlar hakkında doğru ölçümler veren araçları

ŞEKİL 2.3 Psikolojide kullanılan araçların doğrulukları ve hassasiyetleri etkileyici derecede gelişti.





gerektirir. Anketler ve testler, psikologların davranışı değerlendirmek için kullandıkları popüler araçlardır. Gözlemcilerin kullandıkları dereceleme ölçekleri de böyledir. Örneğin, Çocuklarda saldırganlığı derecelemek için kullanılan 7 dereceli bir saldırganlık ölçeğinin bir ucunda "hiç saldırgan değil" (1) yer alırken, diğer ucunda ise "çok saldırgan" (7) yer alır; böyle bir ölçek saldırganlığın, kesin olmamakla birlikte, nispeten doğru bir ölçümünü yapabilir. Olabildiğince doğru ve kesin ölçme araçları kullanmak, davranış alanında çalışan bilim insanlarının sorumluluğundadır.

Ölçme

Bilim insanları, bilimsel yöntemin özelliği olan kontrollü gözlemleri kaydetmek için, iki tip ölçüm kullanırlar. Bu bilimsel ölçümlerden birisi *fiziksel ölçüm*dür; bu ölçüm, üzerinde anlaşma sağlanmış olan bir standart ve ölçme aracı olan boyutları içerir. Örneğin uzunluk, üzerinde uzunluk birimlerinin standartları (örneğin, inç, metre gibi) ve ölçme araçları (örneğin, cetvel) konusunda anlaşma sağlanmış olan bir boyuttur. Ağırlık ve zaman birimleri de aynı şekilde fiziksel ölçümlerdir.

Ama psikoloji alanındaki araştırmaların çoğu, fiziksel ölçümler kullanmaz. Güzelliği, saldırganlığı veya zekâyı ölçmeye yarayacak bir cetvel yoktur. Bu gibi boyutlar, ikinci tip bir ölçmeyi gerektirir –*psikolojik ölçme*: Bir anlamda, insan psikolojik ölçmenin aracıdır. Daha açık ifade edecek olursak, *araştırmacılar arasında sağlanan uyum, psikolojik ölçmenin temelidir*. Örneğin, eğer çeşitli bağımsız araştırmacılar belirli bir eylemin 7 dereceli bir saldırganlık ölçeği üzerinde 3 derecesinin karşılığı olduğu üzerinde anlaşırlarsa, bu, saldırganlık eyleminin psikolojik bir ölçümü olur.

Temel Kavram

Ölçümlerin geçerli ve güvenilir olması gerekir. Genel olarak **geçerlik**, bir ölçmenin "doğruluğunu" gösterir. Bir kavramın geçerli bir ölçümü, ölçtüğünü iddia ettiği şeyi ölçen bir ölçümdür. Varsayın ki bir araştırmacı zekâyı, bir topu burnunun üzerinde dengede tutma süresi olarak tanımlıyor. "İşevurukluk" prensibine göre bu tanım son derece geçerli bir tanımdır. Ama birçoğumuz bu dengede tutma eyleminin, zekânın geçerli bir ölçütü olduğundan kuşkuya kapılırız. Bu ölçümün geçerliği ne zaman desteklenir? İnsanlar zekâyı ölçtüğü kabul edilen diğer görevlerde de dengede tutma görevindeki kadar başarılı olurlarsa, dengede tutma görevinin zekâyı ölçtüğü düşüncesi desteklenir.

Temel Kavram

Bir ölçümün **güvenirliği**, o ölçümün tutarlılığını gösterir. Çeşitli güvenirlik ölçütleri vardır. Bir aracın güvenirliğinden söz ettiğimiz zaman, o aracın düzenli bir şekilde çalışıp çalışmadığından söz etmiş oluyoruz. Bazen çalışan bazen de çalışmayan bir araba çok güvenilir değildir. İki veya daha fazla bağımsız gözlemci yaptıkları gözlem üzerinde anlaşma sağlıyorlarsa, bu gözlemlerin güvenilir olduğu söylenir –yani, bir gözlemcinin gözlemi diğer(ler)inin gözlemi ile uyumlu ise. Örneğin, araştırmacılar 50 üniversite öğrencisinin, bir başka üniversiteden öğrencilerin fotoğraflarını incelemelerini ve bu öğrencilerin ne kadar güvenilir olduklarını 7 dereceli bir ölçek üzerinde işaretlemelerini istediler (1 = hiç güvenilir değil, 7 = çok güvenilir). Sonuçlar, dereceleme yapan öğrenciler arasında yüksek bir uyum olduğunu gösterdi; bu, dürüstlük niteliğinin güvenilir bir şekilde ölçülebileceğini göstermektedir (Rule, Krendl, Ivcevic ve Ambady, 2013). İlginç şekilde, fotoğrafları incelenen (diğer üniversitedeki) öğrencilere de bir dizi araştırma görevi verildi; bu görevlerden birisi de standart zamandan daha fazla zaman verilerek hile yapmalarına fırsat tanıyan bir testti. Öğrencilerin %57'si, bu testte hile yaptı. Araştırmacılar güvenirlik derecelemesinin, öğrencilerin hile yapmaları ya da yapmamalarıyla tamamen ilişkisiz olduğunu buldular. Bu demektir ki, bir ölçümün güvenilir olması, her zaman o ölçümün geçerli olduğu anlamına gelmez.

Ölçümün geçerliği ve güvenirliği, psikoloji araştırmalarında çok önemli bir meseledir. Gelecek bölümlerde farklı araştırma yöntemlerini açıklarken, ölçümlerinin geçerlik ve güvenirliğini belirlemek için araştırmacıların kullandıkları çeşitli teknikleri de açıklayacağız.

Hipotezler

Bir hipotez, bir şeyin geçici açıklamasıdır. Hipotezler çoğunlukla "Nasıl?" ve "Neden?" sorularını cevaplamaya çalışırlar. Bir düzeyde, bir hipotez sadece belirli değişkenlerin ne şekilde ilişkili olduklarına dair bir açıklama öne sürer. Örneğin, psikoloji araştırmalarında yeni bir alan, çevreci ürünler çevreci olmayan ürünlerden daha pahalı, daha gösterişsiz ve daha az etkili olduğu hâlde, insanların neden bunları satın aldığını araştırıyor. Bunun bir örneği, kendisinden daha konforlu ve daha verimli araçlarla aynı fiyata satılan Toyota Prius'tur. Çevreci ürünler satın almayı açıklayabilecek bir *hipotez*, fedakârlık yapmaktır; yani, başkalarına yararlı olmak için özverili davranma eğilimi (Griskevicius, Tybur ve Van den Berg, 2010). Çevreci ürünler satın alma özverili bir davranış olarak görülebilir, çünkü özverili alıcıya daha büyük maliyeti olsa da, çevre ve toplum kazançlı çıkmaktadır.

Yeni teorisyenler "yarışmacı fedakârlık" açıklaması ortaya attılar; bu kavrama göre bireyler özveride bulunuyorlar, çünkü toplumdan yana ve özverili görünmek, toplum içinde daha iyi tanınmalarını sağlamakta ve statülerini yükseltmektedir (Griskevicius ve ark., 2010). Bu nedenle, çevreci ürünler satın alma gibi fedakâr davranışlar, bir kimsenin daha yüksek statüsünün göstergesi olabilir –özverili davranmak için yeterli zamanı, enerjisi, serveti ve diğer kaynakları olan birisi olma olasılığı. Griskevicius ve arkadaşları şöyle bir hipotez kurdular: İnsanların statü arzusunu harekete geçirmek, onları daha lüks ve çevreci olmayan ürünler yerine çevreci ürünler seçmeye yöneltir.

Griskevicius ve arkadaşları (2010) bu hipotezi test etmek için üç deney yaptı. Bu deneylerde araştırmacılar, iki koşul kullanarak, üniversite öğrencilerinin statü isteğini değişimlediler: Statü ve kontrol koşulu. Statü isteğini harekete geçirmek için, bu koşuldaki öğrencilere üniversiteden mezun olma, bir iş arama ve nihayet, yükselme olanağı olan iyi bir şirkette çalışma konulu kısa bir hikâye okutuldu. Kontrol koşulundaki katılımcıların okuduğu hikâye ise kayıp bir konser biletini aramayı, bileti bulmayı ve daha sonra konsere gitmeyi anlatıyordu. Hikâyeyi okuduktan sonra katılımcılar ikinci ve tüketici tercihleri ile ilişkisiz bir görev yapmakta olduklarına inandılar. Muhtemelen satın alabilecekleri şeyleri belirlediler (örneğin araba, bulaşık makinesi, sırt çantası gibi); her bir durumda, daha lüks ve çevreci olmayan ve çevreci olan şeyler birlikte sunuldu. Araştırmacılar, statü yükseltme grubunun, kontrol grubundan daha fazla çevreci malları satın aldıklarını buldular (Deney 1). Üstelik, statü yükseltme grubu sadece alışverişin başkaları tarafından da görüldüğünü düşündüğü zaman çevreci ürünleri satın aldı, kimsenin görmediği (online) durumlarda almadı (Deney 2) ve çevreci ürünler daha pahalı oldukları zaman satın alındı (Deney 3).

Teorik düzeyde, bir hipotez belirli değişkenlerin ilişki şekli hakkında bir açıklama ("neden") öne sürebilir. Griskevicius ve arkadaşları, iki değişken arasında bir ilişki buldular: Statü arzusu ve çevreci ürünler satın alma olasılığı. Yarışmacı fedakârlık teorilerine dayanarak, bu değişkenler ilişkilidir çünkü insanlar özverili davrandıkları zaman (örneğin, çevreci ürünler satın aldıkları zaman) statü satın almaktadırlar. Bu bulgunun pratik bir çıkarımı şudur ki çevreci ürünlerle çevre tahribatı arasında ilişki kurarak veya fiyatlarını düşürerek değil de çevreci ürünlerle yüksek statü arasında bir bağlantı kurulduğu zaman (örneğin, ünlü kişiler onayladıkları zaman) bu ürünlerin satışı artabilir.

Hemen herkes, hayatının bazı dönemlerinde insan davranışını açıklamak için hipotezler öne sürmüştür. İnsanlar neden anlamsız görünen şiddet davranışlarına başvururlar? İnsanlar neden sigara içmeye başlar? Neden bazı öğrenciler okulda diğerlerinden daha başarılıdır? Bilimsel hipotezleri günlük ve rastgele hipotezlerden ayıran bir özellik, *test edilebilirlik*tir (testability). Eğer bir hipotez test edilemezse, bilim için bir değer taşımaz (Marx, 1963). Üç tip hipotez, "test edilebilirlik testi"ni geçemez. Hipotezin kavramları yeterince açık tanımlanmadığı zaman, hipotez kendisini yine kendisiyle tanımladığı zaman (kısır döngü), veya bilimin kabul etmediği etkenleri kullandığı zaman test edilemez.

Hipotezin dayandığı kavramlar yeterince açık değilse ve ölçülemiyorsa, o hipotez test edilemez. Örneğin, ünlü bir kişi vurulduktan sonra neden vurulduğunu açıklamak için "çünkü suikastçının akıl sağlığı bozuktur" demek, "akıl sağlığı bozuk" ifadesinin tanımında anlaşamadığımız sürece, test edilebilir bir hipotez değildir. Ne yazık ki psikologlar ve psikiyatrlar "akıl sağlığı bozuk" gibi terimlerin anlamı üzerinde her zaman anlaşamıyorlar çünkü bu kavramların üzerinde anlaşılmış işevuruk tanımları çoğunlukla yoktur. Bir psikoloji dersinizde, Freud'un birçok hipotezinin test edilebilir olmadığını öğrenmiş olabilirsiniz. Test edilebilir değildir çünkü Freud'un teorilerindeki birçok temel kavramın (örneğin, *id, ego, süperego* gibi) açık işevuruk tanımları ve ölçümleri yoktur.

Hipotez, eğer kendisini kendisiyle tanımlıyorsa, yine test edilebilir değildir. Kendisini kendisiyle tanımlayan bir hipotez demek, bir olayı açıklamak için o olayın kendisini kullanmak demektir (Kimble, 1989, s. 495). Örneğin, "sekiz yaşındaki bir çocuğun dikkati okulda kolayca dağılabilir, çünkü dikkat kusuru vardır" ifadesini ele alın. Bu ifadedeki dikkat kusuru, dikkatini toplayamama ile tanımlanmıştır. Bu nedenle bu ifade basitçe diyor ki çocuk dikkat edemez, çünkü dikkat edemez –yani hipotezin açıklaması bir kısır döngüdür.

Hipotez, bilimin tanımadığı düşüncelere veya etkenlere başvurduğu zaman da test edilebilir değildir. Bilim gözlenebilir, gösterilebilir, görgül olanı ele alır. Korkunç şiddet eylemleri yapanların Şeytan tarafından yönetildiğini öne sürmek test edilebilir bir açıklama değildir, çünkü bilimin alanı içinde olmayan bir güce (Şeytan'a) başvuruyor. Bu gibi hipotezler filozoflar veya din bilimciler için değerli olabilir ama bilim insanları için olmaz.

BİLİMSEL YÖNTEMİN AMAÇLARI

- Bilimsel yöntemin şu dört amacı gerçekleştirmesi beklenir: Betimleme, kestirim, açıklama ve uygulama.

Bu bölümün ilk kısmında, günlük düşünme biçimlerimizin hangi yönlerden bilimsel yöntemden farklı olduğunu ele aldık. Bundan sonraki kısmında ise bilimsel yöntemin amaçlarını inceleyeceğiz. Psikologların bilimsel yöntemi kullanmalarının nedeni, şu dört amaca ulaşmaktır: Betimleme, kestirim, açıklama, ve uygulama (bakınız, Tablo 2.2).

Betimleme

- Psikologlar, olayları ve değişkenler arasındaki ilişkileri betimlemeyi amaçlar; araştırmacılar çoğunlukla nomotetik (nomothetic) yaklaşımı ve nicel analizi kullanırlar.

Betimleme, araştırmacının olayları ve bu olayların ilişkilerini tanımlamak, sınıflamak, listelemek veya kategorileştirmek için kullandığı işlemlere denir. Örneğin, klinik araştırma zihinsel bozuklukları kategorileştirmeleri için uygulamacılara ölçütler sağlar. Bu ölçütlerin çoğu *American Psychiatric Association*'ın yayımladığı *Diognastic and Statistical Manual of Mental Disorders*'da (5'inci baskı, 2013) bulunabilir ki bu kitap DSM-5 olarak da bilinir (bk. Şekil 2.4). Bir örnek olarak, seçici dilsizlik (selective mutism) adı verilen bozukluğu tanımlamak için kullanılan ölçütü ele alın.

TABLO 2.2 PSİKOLOJİK ARAŞTIRMANIN DÖRT AMACI

Amaç	Yapılan İş	Örnek
Betimleme	Araştırmacı zihinsel süreçleri ve davranışı betimlemek için olayları ve olayların ilişkilerini tanımlar, sınıflar, listeler, veya kategorize eder.	Psikologlar depresyonda çaresizliğin, bir eylemi başlatamama ve gelecek hakkındaki kötümserlik gibi belirtilerini betimler.
Kestirim	Araştırmacılar değişkenler arasındaki korelasyonları belirledikleri zaman, zihinsel süreçleri ve davranışları kestirebilirler.	Depresyonun düzeyi arttıkça, bireyler daha fazla çaresizlik belirtileri gösterirler.
Açıklama	Araştırmacılar, neden(ler)i belirleyebildikleri zaman bir olayı açıklayabilirler.	Çözemedikleri problemlerle karşılaşan katılımcılar, çözülebilir. problemlerle karşılaşanlara kıyasla, daha kötümser ve yeni görevleri yapmaya daha isteksiz (yani, çaresiz) hâle gelirler.
Uygulama	Psikologlar bilgilerini ve araştırma yöntemlerini insanların hayatını iyileştirmek için kullanırlar.	Depresyonda olanları üstesinden gelebilecekleri veya kolayca başarabilecekleri görevleri yapmaya teşvik eden tedavi bu insanların çaresizliğini ve kötümserliğini azaltır.

Zechmeister, Zechmeister ve Shaughnessy (2001, s. 12) Tablo 1.2'den alınmıştır..

Seçici Dilsizliğin Tanı Ölçütleri

(a) Başka ortamlarda konuştuğu hâlde, konuşması beklenen özel sosyal ortamlarda (örneğin, okulda) sürekli olarak konuşma başarısızlığı.
(b) Bu rahatsızlık, eğitimde ve işte başarıyı veya sosyal iletişimi engelleyebilir.
(c) Rahatsızlığın süresi en az bir aydır (okulun ilk ayı ile sınırlı değildir).
(d) Konuşma başarısızlığının nedeni, sosyal ortamda konuşulan dile ait bilgi (veya rahatlık) eksikliği değildir.
(e) Rahatsızlık, bir iletişim bozukluğu (örneğin, çocukluktan gelen bir akıcılık kusuru) ile daha iyi açıklanmaz ve özellikle otizm çerçevesindeki bir rahatsızlık, şizofreni veya başka bir psikotik rahatsızlık döneminde ortaya çıkmaz. (DSM-5, 2013, s. 195)

Seçici dilsizliği tanımlamak için kullanılan tanısal ölçütler, bu rahatsızlık için işevuruk bir tanım sağlıyor. Seçici dilsizlik göreceli olarak nadirdir; bu nedenle, bu rahatsızlığı olan bireyler hakkındaki bilgimiz "vaka araştırması"na dayanıyor. Araştırmacılar ayrıca klinikçilere, çeşitli belirtiler ile cinsiyet ve yaş gibi diğer değişkenler arasındaki ilişkilerin yanı sıra, zihinsel hastalıkların yaygınlığına dair betimlemeler sağlamaya çalışıyorlar. Örneğin, araştırmalara göre seçici dilsizlik en çok küçük çocuklar arasında görülüyor ve seçici dilsizlik ile sosyal kaygı arasındaki kuvvetli ilişki yüzünden, seçici dilsizlik sınıflamada şimdi kaygı bozuklukları arasında yer alıyor (DSM-5, 2013). Bölüm 9'da, seçici

ŞEKİL 2.4 Klinikçiler zihinsel hastalıkları, American Psychiatric Association'ın yayımladığı *Diagnostic and Statistical Manual of Mental Disorders*'da (DSM) yer alan ölçütlere göre sınıflar.

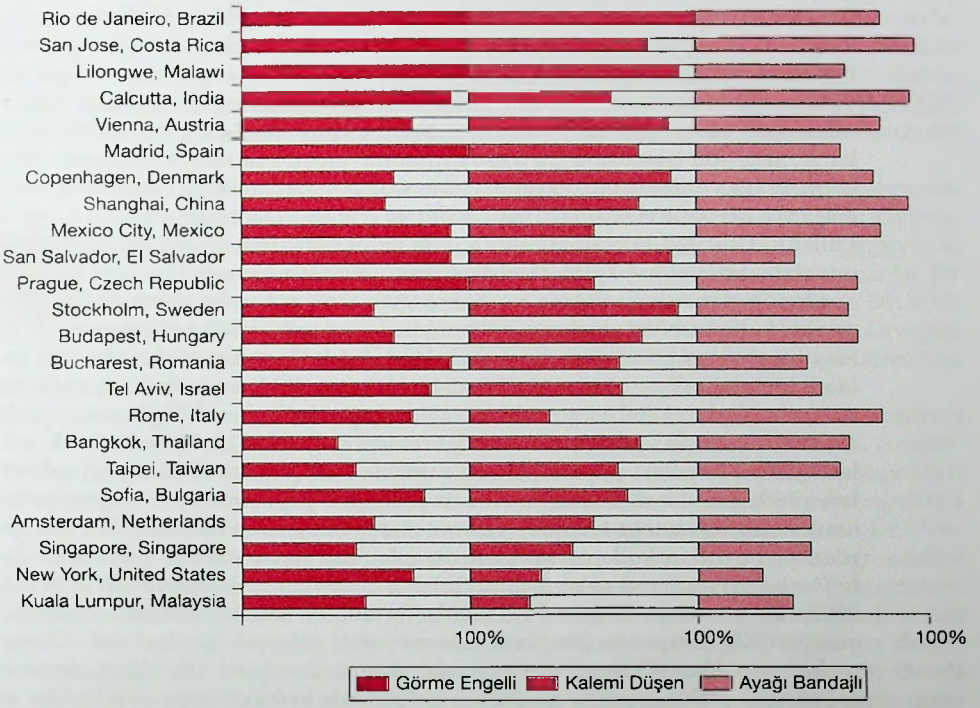


dilsizlik sorunu olan genç bir kızı tedavi etmek için psikologların kullandıkları bir araştırma desenini ele alacağız.

Bilim, psikoloji bilimi dâhil, *nomotetik yaklaşımı* kullanarak olayların betimlemesini geliştirir. "Nomotetik", genel bilimsel yasaların keşfi demektir. Psikologlar davranış hakkında, çeşitli grupları barındıran bir insan evrenine uygulanabilecek geniş genellemelere ve genel yasalara ulaşmak isterler. Bu amaca ulaşmak için, psikoloji araştırmaları çoğunlukla çok sayıda katılımcıyla yapılır. Ve araştırmacılar, bir grubun "ortalama" veya tipik performansını betimlemeye çalışır. Bu ortalama, gruptaki herhangi bir kişinin performansını betimleyebilir de betimlemeyebilir de.

Örneğin, büyük bir kültürler arası araştırmanın bir amacı, farklı ülkelerdeki büyük şehirlerde yardımın derecesini betimlemekti (Levine, Norenzayan, ve Philbrick, 2001). Araştırmacılar, vatandaşların yardım edip etmediklerini belirlemek için, kent merkezinde üç farklı yardım durumu oluşturdular. Her bir ülkede, bir deneyci (1) caddeden karşıya geçmeye çalışan bir görme engelli taklidi yaptı; (2) yolda yürürken kalemini düşürdüğünü fark etmeyen birisini taklit etti; ve (3) ayağında büyük bir bandaj varken, elindeki bir deste dergiyi yere düşürdü ve toplamaya çalıştı. Yirmi üç şehirden 1198 kişi bu üç durumdan birisinde gözlendi. Araştırmanın sonuçları Şekil 2.5'tedir. Ortalama olarak, en çok yardımı Rio de Janeiro'daki insanlar, en az yardımı ise Kuala Lumpur'daki insanlar yaptılar. New York yardım sıralamasında sondan ikinci oldu. Bu bulgular Rio de Janeiro'daki *bütün* insanların yardımsever olduklarını, Kuala Lumpur'daki insanların ise yardımsever olmadığını göstermez. Bu veriler,

ŞEKİL 2.5 Dünyanın çeşitli ülkelerindeki 23 şehirde, üç farklı durumda (görme engelli birisine yardım, yere düşürülen bir kalemi alma, ve bacağını inciten birisine yardım) başkalarına yardım edenlerin oranı; toplam yardım edenlerin sayısına göre yapılmış olan derecelemede Rio de Janeiro en üst, Kuala Lumpur ise en alt sırada yer almıştır. Levine ve arkadaşlarından (2001, Tablo 2) alınmıştır.



genel olarak veya *ortalama olarak*, bu araştırmadaki 23 şehirde insanların yardımseverliğinin farklı olduğunu gösterir.

Nomotetik yaklaşımı kullanan araştırmacılar, bireyler arasındaki büyük farklılıkları önemli bulurlar; bununla birlikte, farklılıklardan çok benzerliklere üzerinde dururlar. Örneğin, bir kimsenin birey olarak özgünlüğü, o kimsenin kalbinin, herkesin kalbi gibi göğsünün sol üst boşluğunda olduğunu bilmemizle tehlikeye düşmez. Aynı şekilde, bir kimsenin davranışlarının pekiştirmeyle (örneğin, ödül, ceza) değişeceğini söylediğimiz zaman da o insanın özgünlüğünü inkâr etmiş olmuyoruz. Araştırmacılar sadece, farklı organizmaların oluşturduğu bir grubun ortalama performansına dayanarak, organizmaların genel olarak nasıl olduklarını betimlemeye çalışırlar.

Başta Gordon Allport (1961) olmak üzere bazı psikologlara göre, nomotetik yaklaşım yeterli değildir; özgün bireyler, ortalama bir değerle betimlenemez. *İdiografik yaklaşımı* kullanan araştırmacılar, gruplardan ziyade, bireyleri incelerler. Bu araştırmacılara göre, her ne kadar bireyler genel yasalara veya prensiplere uygun şekilde davranıyorlar ise de onların özgün yönleri de betimlenmelidir.

İdiografik araştırmaların en önemli şekli, Bölüm 9'da ele alacağımız tek denek araştırmasıdır.

Araştırmacılar, cevaplamak istedikleri araştırma sorularına bağlı olarak, grupları mı yoksa bireylerin davranışlarını mı araştıracaklarına karar verirler. Bazı araştırmacılar bunlardan birisini seçerken, bazıları da her ikisini birden kullanır. Örneğin bir klinik psikolog, tedavi altındaki birkaç danışanı için esas olarak idiografik araştırma yöntemini uygulamaya karar verebilir ama çok sayıda katılımcı gerektiren bir araştırma sorusunu cevaplaması gerektiği zaman nomotetik bir yöntem kullanabilir. Araştırmacının vermesi gereken bir diğer karar, nicel bir araştırma mı yoksa nitel bir araştırma mı yapacağıdır. *Nicel araştırma*da veriler, istatistiksel özet ve analizler kullanılarak betimlenir. *Nitel araştırma* ise verileri sözlü olarak özetler ve çok az istatistiksel özet veya analiz kullanır. Psikoloji araştırmalarında nomotetik araştırma idiografik araştırmadan daha fazla olduğu gibi, nicel araştırma da nitel araştırmadan daha fazladır.

Sosyologlar, antropologlar, ve psikologlar, nitel araştırma yöntemini yaygın şekilde kullanır. Nitel araştırma yönteminin bir tek şekli yoktur (Silverman, 2011) ama bu yaklaşımı kullanan araştırmacılar çoğunlukla anlatımsal bir kayıtta anlamlı konular ve kategoriler ararlar; daha sonra, yaptıkları gözlemlerin sözlü bir özetini sunarlar. Nitel araştırmalar olayları çoğunlukla kendi bağlamları içinde ele alırlar ve kişisel görüşmelere ve davranışın doğrudan gözleminden elde edilen ayrıntılı kayıtlara dayanırlar. Son zamanlardaki analizlerde, internet sohbetleri gibi kaynaklar da kullanılmaktadır. Görüşme yapılırken nitel araştırmacılar katılımcılarından, teorisyenlerin ve önceki araştırmacıların hazırladıkları kategorileri ve boyutları kullanan soruları cevaplamalarını istemekten çok, kendi deneyimlerini *kendilerine* anlamlı gelen şekilde betimlemelerini isteyebilirler (Kidd, 2002). Bu yöntemi Kidd ve Kral (2002) Toronto'da sokakta (29'uncu caddede) yaşayan 17-24 yaş arası gençlerin deneyimlerini anlamak için kullandı. Görüşmelerin odaklandığı hususlardan birisi, intihar deneyimi idi. Görüşülenlerin çoğunluğu (%76) geçmişte de intihar girişiminde bulunduklarını bildirdiler ve anlattıkları analiz edilince görüldü ki intihar girişimi ve deneyimi çoğunlukla yalnızlık, reddedilme/ihanet, değersizlik ve fahişelik duygularıyla yakından ilişkilidir. Önemli bir nokta olarak araştırmacılar bu araştırmanın, sokak çocuklarının intihar deneyimleri ile bağlantılı olan ve daha önceki araştırmaların belirleyemediği birçok hususu belirlediğini bildirdiler. Şöyle ki, "kontrolün kaybı, fahişelik esnasında saldırı, 'yavaş intihar' olarak uyuşturucu kullanımı ve yakın ilişkilerin bozulması", bu gençlerin intihar deneyimleriyle bağlantılı olan etkenlerdir (s. 411). Davranış gözleminin anlatımsal kayıtlarını ele alacağımız Bölüm 4'te, nitel araştırmaların başka örneklerini bulacaksınız; Bölüm 9'da açıklanacak olan vaka araştırmaları da nitel araştırma yönteminin bir şeklidir.

Kestirim

- Korelasyonel ilişkiler, psikologlara davranışı veya olayları kestirme olanağı verir ama bu ilişkiye neyin neden olduğunu söyleme olanağı vermez.

Olayların ve bu olaylar arasındaki ilişkilerin betimlenmesi çoğunlukla kestirim (prediction) için bir temel sağlar ki bu, bilimsel yöntemin ikinci amacıdır. Psikolojideki birçok önemli soru, kestirim yapmayı gerektirmektedir. Örneğin,

ebeveynini küçük yaşta kaybeden bir çocuğun depresyona eğilimi daha fazla mı olur? Çok saldırgan olan çocukların yetişkinlikte duygusal sorun yaşama olasılıkları daha mı fazladır? Stres, daha çok fiziksel rahatsızlığa yol açar mı? Araştırma bulgularına göre tüm bu soruların cevabı "evet"tir. Bu bilgiler psikoloji bilimini sadece zenginleştirmekle kalmaz, fakat aynı zamanda hem tedavide hem de duygu durum bozukluklarının önlenmesinde yararlıdır. Bundan başka, birçok psikoloğun önemli bir uğraşı alanı psikolojik testlerdir; bu alanda testler, bireylerin gelecekteki davranışını ve performansını (örneğin, işteki performansını) kestirmede kullanılır. Siz bunların bazılarına zaten aşinasınız; örneğin üniversitelere ve meslek okullarına giriş testleri böyledir.

Temel Kavramı

Bir değişkenden alınan puanlar, ikinci bir değişkenin puanlarını kestirmede kullanılabilir; bu durumda biz, bu iki değişkenin korelasyonlu olduğunu söyleriz. Aynı insanlardan, olaylardan veya şeylerden alınan iki ölçüm birlikte değişiyorsa, orada bir **korelasyon** vardır –yani, eğer belirli bir değişkenden alınan puanlar, bir başka değişkenden alınan belirli puanlarla bağlantılı olmaya eğilimliyse. Böyle bir bağlantı olduğu zaman, puanların "birlikte değiştiği" (covary) söylenir. Örneğin, stres ile hastalığın korelasyonlu olduğu biliniyor; insanlar ne kadar çok stresli bir hayat sürerlerse, o kadar daha fazla fiziksel hastalığa yakalanırlar.

Muhtemelen deneyim sahibi olduğunuz bir ölçmeyi ele alın; bu deneyim örneğin, sınıflarda yaptığınız öğretmen/ders değerlendirmesi olsun. Araştırmacılar, derse kayıtlı *olmayan* öğrencilerin öğretmenler hakkında yaptıkları değerlendirmelerin, derse kayıtlı *olan* öğrencilerin dönem sonunda yaptıkları değerlendirmelerle ne ölçüde korelasyonlu olduğunu öğrenmek istediler (Ambady ve Rosenthal, 1993). Araştırmacılar bir grup kadın üniversite öğrencisine üç farklı araştırmada 30 saniye, 10 saniye, ve sadece 6 saniye süreyle öğretmenlerin sessiz video kliplerini gösterdiler. Araştırmacılar, bu "kısa süreli sessiz davranış"a dayanan değerlendirmelerin, dönem sonundaki öğretmen değerlendirmesi ile korelasyonlu olduğunu buldular. Yani, öğretmenlerin derste daha olumlu değerlendirilmesi, onların videodaki davranışlarından daha yüksek puan almalarıyla bağlantılı oldu; aynı şekilde, derste daha olumsuz değerlendirilmeleri, videodaki davranışlarından daha düşük puan almalarıyla bağlantılı oldu. Bu nedenle biz, öğretmenlerin dersteki beğenilme düzeylerini, videodaki kısa süreli davranışlarına bakarak kestirebiliriz. Bu sonuçlar gösteriyor ki gözlemciler, etkili davranışları (örneğin, beğenilme gibi) çok çabuk ve oldukça doğru bir şekilde değerlendirebilirler.

Şu hususa özellikle dikkat edilmelidir ki başarılı bir kestirim her zaman, iki değişken arasında *neden* bir ilişki olduğunu bilmeyi gerektirmez. Bazı insanların, hayvanların davranışlarını gözlemenin depremi önceden bilmeye yardımcı olduğunu söylemeleri olgusunu ele alın. Görünüşe bakılırsa bazı hayvanlar, depremden hemen önce alışılmadık şekilde davranırlar. Bu nedenle havlayan ve daireler çizerek koşan bir köpek ve deliğinden kaçan bir yılan, depremin güvenilir bir habercisi olabilir. Eğer öyleyse, bu hayvanlar insanları yaklaşan felaketlere karşı uyarmak için kullanılabilir. Hatta, deprem olasılığı yüksek alanlarda yaşayan insanların, kendilerini uyarmaları için belirli hayvanları gözlemeleri istenebilir; tıpkı bir zamanlar madencilerin dehlizlerde kanarya tutmaları gibi. Bunu yapmak için, bazı hayvanların depremden önce *neden* tuhaf şekilde davrandıklarını veya hatta depremin neden meydana geldiğini

bilmemiz bile gerekmez. Üstelik, biz hiçbir zaman bir hayvanın tuhaf bir davranışının depreme neden olduğunu öne sürmüyoruz.

Yardım davranışındaki kültürlerarası farklar üzerine yaptıkları araştırmalarında Levine ve arkadaşları'na (2001) göre bir şehirdeki yardımseverliğin derecesini, insanların satın alma gücü ölçütüne dayanan ekonomik refah göstergesine bakarak kestirebiliriz. Bir kentte yaşayanların satın alma gücü *düşük* ise yardım etme olasılığı daha yüksektir. Bu araştırmacılar düşük satın alma gücünün yardımseverliğe *neden olduğunu* iddia etmediler; bir ülkenin ekonomisi zayıf olduğu zaman, geleneksel değerler sisteminin yardımseverlik davranışını teşvik ettiğini öne sürdüler.

Bir başka araştırmacı grubu ise yardım davranışında gözlenen bu kültürel farkların nedeninin geleneksel değerler olabileceğini düşündü ve "kuşatıcı kültürler" (embedded culture) kavramını araştırdı (Knafo, Schwartz ve Levine, 2009). Kuşatıcı kültür demek, bireylerin kendilerini aileleriyle veya içinde yaşadıkları grupla özdeşleştirerek hayata anlam verdikleri kültür demektir. Böyle bir kültürde yaşayan bireyler grup üyelerinin mutluluğunu önde tutarlar ve grup dışındaki insanlarla (örneğin, ayağı yaralı veya kalemini düşüren yabancılar gibi) az ilgilenirler. Bu araştırmacılar, daha önceki araştırmada yardım davranışının gözlendiği 23 şehirde kültürel kuşatıcılığı ölçtüler. Beklendiği gibi, bir kültür ne kadar kuşatıcı ise insanların yabancılara yardım etme olasılığı o kadar az çıktı; böylece hipotez desteklenmiş oldu. Her ne kadar bu korelasyonel bulgu geleneksel, kuşatıcı değerlerin insanların yardım davranışını etkileyeceği açıklamasını destekliyor ise de, bu değerlerin yardım davranışındaki farkların *nedeni* olduğunu söyleyebilmek için ilave kanıtlara ihtiyaç vardır.

Açıklama

- Nedensel çıkarım için gerekli olan şu üç koşul karşılandığı zaman, psikologlar bir olayın nedenini anlar: Birlikte değişim, zaman-sıra ilişkisi, ve mümkün alternatif nedenlerin elenmesi.
- Araştırmacıların bağımlı değişken üzerindeki etkilerini belirlemek için bağımsız değişkeni değişimledikleri deneysel yöntem, zaman-sıra ilişkisi kurar ve birlikte değişimi açık bir şekilde belirleme olanağı sağlar.
- Eğer bir araştırmada karıştırıcı faktörler yoksa, bir ilişkinin mümkün alternatif nedenleri bertaraf edilmiş olur.
- Farklı evrenleri, ortamları ve koşulları betimlemek için araştırmacılar, bir araştırmanın bulgularını genelleştirmek isterler.

Betimleme ve kestirim bilimde önemli amaçlar olmakla birlikte, bir olayı açıklayabilmek ve anlayabilmek için bunlar sadece ilk adımlardır. Açıklama, bilimsel yöntemin üçüncü adımıdır. Bir olayın nedenlerini belirleyebildiğimiz zaman, o olayı anlayabilir ve açıklayabiliriz. Araştırmacılar tipik olarak, bir olayın nedenlerini belirlemek için *deneyler* yaparlar. Birçok insan genellikle araştırma hakkında konuşurken "deney" sözcüğünü kullanır ama bu yanlıştır. Deneysel araştırma, betimsel ve korelasyonel araştırmalardan farklıdır çünkü bilim insanı deneylerde üst düzeyde kontrol sağlamak ister. Araştırmacı bir durumu kontrol ederken, bağımlı değişken (ilgilenilen olay) üzerindeki etkisini

belirlemek için, bağımsız değişkenleri birer birer değişimler. Psikologlar kontrollü deneyler yaparak ve nedensel çıkarım yaparak bir olayın nedenine ulaşırlar. Psikologlar için deneyler, nedensel çıkarım yapabilmeleri için çok önemlidir; bu nedenle biz, deneysel yöntemi ayrıntılı şekilde açıklamak için, Bölüm 6, 7, ve 8'i deneysel yönteme ayırdık.

Temel Kavram

Bilim insanı, bir **nedensel çıkarım** yapabilmek için üç önemli koşulu sağlar: *Olayların birlikte değişimi, zaman-sıra ilişkisi* ve *mümkün alternatif nedenlerin elenmesi*. Bu üç koşulu anlamanız için basit bir örnek verelim. Varsayın ki başınızı bir kapıya çarptınız ve başınız ağrıyor; muhtemelen siz, başınızın ağrımasının *nedeninin*, kapıya çarpmanız olduğu çıkarımını yaparsınız. Nedensel çıkarımın ilk koşulu, olayların birlikte değişmesidir. Eğer bir olay bir diğer olayın nedeni ise, iki olayın birlikte değişmesi gerekir; yani, birisi değiştiği zaman, diğerinin de değişmesi gerekir. Bizim örneğimizde kafanızın dik pozisyondan kapıya çarpma pozisyonuna geçmesi ile ağrımıyorken ağrımaya başlaması deneyiminin birlikte değişmesi gerekir.

Nedensel bir çıkarım yapabilmek için sağlanması gereken ikinci koşul, *zaman-sıra ilişkisi*dir. Varsayılan nedenin (başınızı çarpmanız) varsayılan sonuçtan (baş ağrısı) önce gelmesi gerekir. Eğer başınız kafanızı vurmadan önce ağrırsa, başınızı çarptığınız için ağrıdığı sonucuna varamazsınız. Bir başka ifadeyle, baş ağrısı, sizin önce başınızı vurmanıza bağlıdır. Nihayet, nedensel ilişki sadece, diğer muhtemel nedenlerin etkili olmadığı gösterildikten sonra kabul edilir –*mümkün alternatif nedenler bertaraf edildiği zaman*. Bizim örneğimizde bunun anlamı şudur: Başınızı vurduğunuz için başınızın ağrıdığı sonucuna varmanız için, başınızın ağrımasına neden olabilecek diğer mümkün nedenler üzerinde düşünmeniz ve bu nedenlerin etkili olmadığını göstermeniz gerekir (örneğin, zor bir kitap okumak gibi).

Ne yazık ki insanlar, sadece ilk koşul karşılandığı zaman her üç koşulun da karşılandığını düşünme eğilimindedir. Örneğin, sert bir disiplin ve fiziksel ceza kullanan ebeveynlerin çocuklarının, yumuşak bir disiplin kullanan aile çocuklarından daha saldırgan oldukları öne sürülmüştür. Ebeveynlerin disiplini ile çocukların saldırganlığı açıkça birlikte değişiyor. Üstelik, bizim ebeveynin çocuğun nasıl davranacağını etkilediğini varsaymamız olgusu bizi, zaman-sıra koşulunun karşılandığını düşünmeye yöneltebilir –ebeveynler fiziksel ceza uygular ve bu ceza çocuğun saldırganlığı ile sonuçlanır. Ama şöyle bir durum da var: Küçük çocuklar farklı düzeylerde hareketlidirler ve saldırgandırlar ve çocukların davranışları, ebeveynlerin onları kontrol etme tepkilerini çok etkiler. Bir diğer ifadeyle, sert disiplinin saldırganlığa neden olmasından ziyade, bazı çocuklar doğal olarak saldırgan olabilir ve bunları kontrol etmek için sert bir disiplin gerekiyor olabilir. Bu nedenle, nedensel ilişkinin yönü, bizim ilk bakışta düşündüğümüzün tersine olabilir.

Nedensel bir çıkarım yapmanın kilit unsuru, sonuçları iki veya daha fazla koşulda karşılaştırmaktır. Örneğin, varsayın ki bir öğretmen sınıfta bireysel tepki aygıtı ("clickers") kullanmanın, öğrenmeyi desteklediğini göstermek istiyor. Sınıfta, bu sistemle cevaplanması gereken sorular sorabilir ve bu ders tekniğinden sonra öğrencilerin test performansını betimleyebilir. Ama bu noktada,

öğretmen ne biliyor? Başka bir teknikle öğrenen bir diğer grup öğrenci aynı miktarda öğrenmiş olabilir. Öğretmen, gözlediği performansın *nedeni*nin bireysel tepki aygıtı olduğunu öne sürmeden önce, bu yöntemi makul başka yöntemlerle karşılaştırması gerekir. Yani, bireysel tepki aygıtı kullanan grubun öğrenmesi ile bunu kullanmayan grubun öğrenmesi arasında fark olduğunu bulması gerekir. Böyle bir bulgu, öğretme stratejisi ile performansın birlikte değiştiğini gösterir. Kontrollü bir deney yapılırken bağımsız değişken ile bağımlı değişken birlikte değişiyorsa, böyle bir sonuç bir "ödül" olarak görülür. Nedensel bir çıkarım için zaman-sıra koşulu karşılanır, çünkü araştırmacı bağımsız değişkeni (örneğin, öğretme tekniğini) değişimler ve *daha sonra* her iki durumdaki (bağımsız değişkenin uygulandığı ve uygulanmadığı durumlardaki) bağımlı değişken düzeyleri (öğrenme düzeyleri) arasındaki farkı ölçer.

Araştırmacının nedensel bir ilişki kurabilmesi için karşılaması gereken üç koşulun en zoru, mümkün diğer nedenleri bertaraf etmesidir. İki farklı öğretim yaklaşımının (bireysel tepki sistemi var/yok) etkisini değerlendiren bir araştırmayı ele alın. Araştırmacının katılımcıları dersin sabah bölümünde kişisel tepki sistemi kullanarak, öğleden sonraki koşuluna ise kullanmadan öğretim koşullarına atadığını varsayın. Eğer böyle yapılsaydı, iki grup arasında öğrenme düzeyinde bulunan bir farkın nedeni öğretim yöntemi farkı olabilirdi *veya* öğleden önce/öğleden sonra farkı da olabilirdi. Bu nedenle araştırmacı, iki grup arasındaki performans farkının nedeninin test etmiş olduğu bağımsız değişken mi (bireysel tepki sisteminin olması ve olmaması mı) yoksa iki grup arasındaki öğleden önce/öğleden sonra farkı mı olduğunu belirleyemez. Daha formel bir ifadeyle, öğretim yöntemi bağımsız değişkeni ile uygulama zamanı bağımsız değişkeni "karışır". Etkili olma potansiyeli olan iki değişkenin eş zamanlı olarak birlikte değişmesine izin verildiği zaman **karışma** ortaya çıkar. Araştırmada karışma ortaya çıktığı zaman, performansta gözlenen farkın nedeninin hangi değişken olduğunu belirlemek olanaksızdır.

Temel Kavram

Araştırmacılar, olayların nedenlerini açıklamak için deneyler yaparlar. Böyle olmakla birlikte, dikkatle düzenlenmiş bir deney araştırmacının nedensel bir çıkarım yapmasına izin verdiği zaman bile, cevaplanması gereken birçok soru kalır. Önemli bir soru, deneyin bulgularının, sadece deneye katılanlar için geçerli olması derecesiyle ilişkilidir. Araştırmacılar çoğunlukla bulgularını, deneye katılmayanları betimlemek amacıyla, genellemek isterler.

Psikoloji araştırmalarına katılanların büyük kısmı, üniversitelerdeki psikolojiye giriş dersi öğrencileridir. Psikologlar sadece üniversiteye yeni başlayanlara uygulanabilecek prensipler mi keşfederler? Benzer şekilde, laboratuvar araştırmaları gerçek ortamlarda olandan çok daha kontrollü koşullarda yapılır. Bu nedenle bilim insanının önemli bir görevi, laboratuvar bulgularının "gerçek hayat"a genelleştirilip genelleştirilemeyeceğini belirlemektir. Bazı insanlar otomatik olarak, laboratuvar araştırmalarının gerçek dünya ile ilişkisiz olduğunu veya gerçek dünyada yararsız olduğunu zanneder. Ama bu kitapta araştırma yöntemlerini öğrendikçe göreceğiz ki laboratuvar bilimi ile gerçek dünya arasındaki ilişkiye dair bu görüşler yararlı ve tatmin edici değildir. Böyle düşünmek yerine psikologlar, her ikisinin de önemli olduğunu kabul ederler: Laboratuvarda elde edilen bulgular gerçek dünyadaki olayları açıklamaya

yardım eder, ve bu bilgiler araştırmalardaki ve müdahalelerdeki gerçek dünya problemlerine uygulanır.

Uygulama

- Uygulamalı araştırmalarda psikologlar bilgilerini ve araştırma yöntemlerini, insanların hayatını iyileştirmek için kullanırlar.
- Psikologlar davranış ve zihinsel süreçler hakkında bilgi edinmek ve teorileri test etmek için temel araştırma yaparlar.

Psikolojide araştırma yapmanın dördüncü amacı, uygulamadır. Psikologlar davranışı ve zihinsel süreçleri betimleme, kestirme, ve açıklama ile uğraşırlar ama bu bilgi boşlukta değildir. Bu bilgi gerçek bir dünyanın içinde yer almaktadır: Bu dünya (karşılaştığımız problemlerden sadece birkaçını söylemek gerekirse) zihinsel hastalıkların olduğu, insanların şiddet ve saldırganlığa kurban gittiği, peşin hükümlerin ve şablonların (stereotype) insanların toplum içinde nasıl yaşayacaklarını ve işlevlerinin ne olacağını etkilediği bir dünyadır. Dünyamızdaki problemlerin listesi, bazen sonsuz gibi görünür ama bu bizim cesaretimizi kırmamalıdır. Psikologların araştırma sorularının ve bulgularının genişliği, araştırmacıların hayatımızın önemli yönlerini ele almalarına yardım etmek ve bireylerin hayatında değişmeler yaratmak için birçok yol sağlar.

Temel Kavram

Değişme yaratmaya yönelik çalışmalara çoğunlukla "uygulamalı araştırmalar" denilir. **Uygulamalı araştırma**larda psikologlar, insanların hayatlarını iyileştirmek için araştırma yaparlar. Zihinsel rahatsızlığı olan insanlar için bu değişme, tedavi teknikleri konusunda yapılan araştırmalarla meydana gelir. Ama uygulama alanında çalışan psikologlar, çok farklı alanlarda çalışırlar; örneğin, okullarda öğrencilerin, işyerlerinde çalışanların, ve toplumda bireylerin hayatını iyileştirmek için çalışırlar.

Temel Kavram

Öte yandan, **temel araştırma** yapan araştırmacılar esas olarak davranışı ve zihinsel süreçleri anlamaya çalışırlar. İnsanlar temel araştırmaları genellikle "bilim için bilim yapmak" olarak ifade ederler. Temel araştırma tipik şekilde, bir olayı açıklamak için öne sürülmüş olan bir teoriyi test etmek amacıyla, bir laboratuvar ortamında yapılır.

Psikolojinin tarihi boyunca temel araştırma ile uygulamalı araştırma arasında bir gerilim olmuştur. Bununla birlikte, geçmiş on yıllar içinde araştırmacılar, psikoloji prensiplerinin insan hayatını iyileştirmesi için önemli, yaratıcı uygulamalarına daha fazla odaklandılar (Zimbardo, 2004). Gerçekte, psikolojinin iyi bilinen prensiplerinin –ki bunlar temel psikoloji araştırmalarında keşfedilmişlerdir– uygulamaları şimdi çok yaygındır; öyle ki insanlar şimdi olağan gördüğümüz şeylerden önceki laboratuvar araştırmalarını unutma eğilimindedirler. Örneğin, pozitif pekiştirme tekniklerinin kullanımı, psikolojik testler ve tedaviler, ve kendini geliştirme günlük hayatın bir parçası hâline geldi. Ayrıca, psikoloji prensiplerinin uygulamaları eğitimde, sağlıkta ve adlî suç ortamında yaygındır. Psikolojinin günlük hayattaki uygulamalarının birçok örneğini şu sitede görebilirsiniz: www.apa.org/research/action.

Önemli bir faktör, temel ve uygulamalı araştırmaları birbirine bağlar: Araştırmalara ve gerçek hayatta uygulamaya öncülük etmesi için teorilerin kullanılması. Gelecek bölümde, psikoloji teorilerinin nasıl geliştirildiğini göreceğiz.

BİLİMSEL TEORİLERİN KURULMASI VE TEST EDİLMESİ

- Teoriler, olayların nedenlerini açıklamak için öne sürülen önerilerdir, ve ait oldukları alana ve açıklama düzeylerine göre farklılıklar gösterirler.
- Bilimsel bir teori, mantıksal olarak organize edilmiş olan bir önermeler dizisidir; bu önermeler olayları tanımlar, olaylar arasındaki ilişkileri betimler ve olayların nasıl meydana geldiklerini açıklar.
- Aracı değişkenler, bağımsız ve bağımlı değişkenlerin neden ilişkili olduklarını açıklamak için, teorilerde kullanılan kavramlardır.
- Başarılı bilimsel teoriler görgül bilgileri organize eder, test edilebilir hipotezler sağlayarak araştırmalara ve titiz test işlemlerine öncülük eder.
- Araştırmacılar teorileri şu yöntemlerle değerlendirir: Teorinin iç tutarlılığını yargılayarak, teori test edildiği zaman beklenen sonuçların çıkıp çıkmadığını gözleyerek ve teorinin özlü açıklamalara dayanan dakik kestirimler yapıp yapamadığına dikkat ederek.

Teoriler, doğanın nasıl işlediğine dair "düşünce"lerdir. Psikologlar, hem insanların ve hayvanların neden öyle davranıp düşündükleri hakkında hem de davranış ve zihinsel süreçlerin niteliği hakkında teoriler öne sürerler. Bir psikoloji teorisi, farklı açıklama düzeyleri kullanılarak geliştirilebilir; örneğin, teori fizyolojik ya da kavramsal bir düzeyde geliştirilebilir (bk., Anderson, 1990; Simon, 1992). Fizyolojik düzeydeki bir şizofreni teorisi, şizofreniyi biyolojik nedenlerle açıklar; örneğin, özel genetik taşıyıcılar gibi. Kavramsal düzeyde geliştirilmiş bir teori ise psikolojik nedenler öne sürer; örneğin, duygusal çatışma veya stres örüntüleri (pattern) gibi. Bir şizofreni teorisinin hem biyolojik hem de psikolojik nedenler içermesi de mümkündür.

Teoriler, alanları açısından farklılık gösterirler –açıklamaya çalıştıkları olayların genişliği açısından. Geniş bir alanı kapsayan teoriler, karmaşık olayları açıklamaya çalışırlar; örneğin, sevgi veya insan bilişi gibi. Genel olarak, bir teori ne kadar geniş bir alanı kapsarsa o kadar da karmaşık olma olasılığı vardır. Böyle olmakla birlikte günümüz psikolojisindeki çoğu teoriler, sınırlı bir olaylar grubunu açıklamaya çalışan göreceli olarak daha küçük alanları kapsayan teorilerdir. Örneğin, "flaş bellek" (flashbulb memory) beklenmedik ve duygusal olayların eşlik ettiği çok özel detayları olan kişisel anıları açıklamaya çalışır; örneğin, 11 Eylül 2001'de meydana gelen korkunç olaya ait anılar böyle anılardır (Brown ve Kulik, 1977).

Bilim insanları teorileri geliştirmek için sezgi, kişisel gözlem, olgu ve düşünceleri birlikte kullanırlar. Ünlü bilim filozofu Karl Popper'a (1976) göre gerçekten yaratıcı teoriler bir probleme duyulan yoğun ilgi ve eleştirel bir hayal gücü ile geliştirilir –eleştirel ve "alışılmışın dışında" düşünme yeteneği. Araştırmacılar, bir problem veya bir araştırma sorusu hakkında nelerin bilindiğini inceleyerek ve aynı zamanda, hataları veya noksanları arayarak bir teori inşa etmeye başlar. Teori kurma süreci, Bölüm 1'de betimlediğimiz araştırmaya başlama ve hipotez kurma sürecine benzer.

Temel Kavram

Her ne kadar teoriler açıklama düzeylerine ve alanlarına göre farklılık gösterir ise de, yine de hepsinin ortak özellikleri vardır (bk. Tablo 2.3). Biz, bilimsel bir **teori** için şu tanımı öneriyoruz: *Mantıkî olarak organize edilmiş önermeler*

TABLO 2.3 TEORİLERİN ÖZELLİKLERİ

Tanım	Bir teori, olayları tanımlamaya, bu olaylar arasındaki ilişkileri betimlemeye, ve bu olayların meydana gelişini açıklamaya yarayan mantıkî olarak organize edilmiş bir önermeler dizisidir.
Alan	Teoriler, açıklamaya çalıştıkları olayların genişliği bakımından farklılık gösterirler; çok spesifik bir olaydan (örneğin, flaş bellek gibi) karmaşık bir olaya (örneğin, sevgi gibi).
İşlev	Bir teori, önceki araştırmaların elde ettiği görgül bilgileri organize eder ve test edilebilir hipotezler önererek yeni araştırmalar yapılmasının önünü açar
Önem	Aracı (intervening) değişkenler, değişkenler arasında açıklayıcı bir bağ kurarlar.
Gelecek	İyi bir teori: • *Mantıklıdır*: Anlamlıdır ve mantıklı olarak çıkarılabilen kestirimlere izin verir. • *Dakiktir*: Davranış hakkında yapılan kestirimler genel değil, spesifiktir. • *Ekonomiktir*: Bir olayın en iyi açıklaması, en basit olanıdır.

Zechmeister, Zechmeister ve Shaughnessy'den (2001, s. 29) alınmıştır.

(iddialar, ifadeler) seti: bu önermelerin işlevi olayları (kavramları) tanımlamak, bu olaylar arasındaki ilişkileri betimlemek ve bu olayların ortaya çıkışını açıklamaktır. Örneğin, bir flaş bellek teorisinin bir flaş anının ne olduğunu ve bir flaş anının diğer anılardan nasıl bir farkının olduğunu kesin şekilde açıklaması gerekir. Bir teori ilişkileri betimler; örneğin, duyguların işe karışma derecesi ile hatırlama miktarı arasındaki ilişki gibi (örneğin, Talarico ve Moore, 2012). Nihayet bir teorinin aynı zamanda, bir kimse flaş anılarının doğruluğundan çok emin olsa bile, bazen bu anıların açıkça yanlış olmasını da açıklamak zorundadır (bakınız, Neisser ve Harsch, 1992). Olaydan bir ve iki yıl sonra insanların 11 Eylül saldırısı hakkındaki anılarını araştıran bir çalışmada olan şey buydu (Conway, Skitka, Hemmerich ve Kershaw, 2009). Zamanla anılarının doğruluğundaki ve tutarlılığındaki düşüşe rağmen katılımcılar, flaş anılarının doğruluğuna olan yüksek güvenlerini korudular.

Bir teorinin başlıca işlevleri, görgül bilgileri *organize* etmek ve yeni araştırmalara öncülük etmektir. Flaş bellek gibi göreceli olarak spesifik araştırma alanlarında bile çok sayıda araştırma yapıldı. Bir araştırma alanı genişledikçe, bu alandaki araştırma sayısı da artar. Bilimsel teoriler önemlidir, çünkü birçok araştırma bulgusunun organize edilmesine ve bulgular arasındaki ilişkilerin belirlenmesine olanak sağlar. Bulguların bu mantıkî organizasyonu araştırmacılara rehberlik eder, çünkü gelecekte yapacakları araştırmalar için test edilebilir hipotezler kurabilirler.

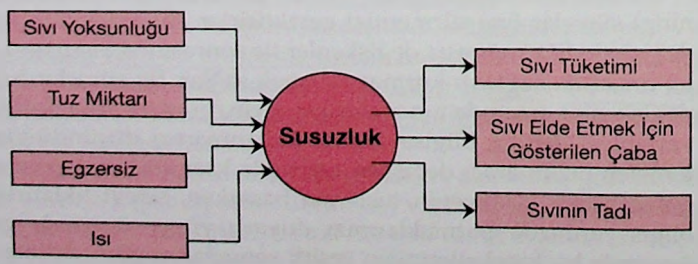
Teoriler sıklıkla, gözlenen davranışları açıklayabilmemiz için aracı (intervening) süreçler öne sürmemizi gerektirirler. Bu aracı süreçler, araştırmacıların değişimlediği bağımsız değişkenler ile sonrasında ölçtükleri bağımlı değişkenler arasında bağlantı kurmaya olanak sağlar. Bu süreçler bağımsız ve bağımlı değişkenler arasında aracı oldukları için, bunlara *aracı değişkenler* (intervening variables) denilir. Bilgisayarınızı kullanmanızı düşündüğünüz zaman, muhtemelen bizim aracı değişken terimiyle kast ettiğimiz şeye aşina olduğunuzu göreceksiniz. Klavyenin tuşlarına basarken, fareyi tıklatırken veya diz üstü bilgisayarınızda parmaklarınızı duyarlı yüzey üzerinde gezdirirken, ekran üzerinde bu hareketlerinizin çeşitli sonuçlarını göreceksiniz. Ama yine de bu

sonuçlara doğrudan neden olan şey, sizin tuşlara, fareye, veya duyarlı yüzeye dokunmanız değildir; buradaki aracı değişken, sizin hareketleriniz ile ekranda gözlediğiniz sonuçlar arasında bağlantı kuran yazılım programıdır.

Aracı değişkenler, bilgisayar uygulamaları gibidir. Sizin tuşlara dokunuşunuz ile ekranda gördükleriniz arasındaki bağlantıya karşılık gelen bir bağlantıyı, bağımsız ve bağımlı değişkenler arasında kurar aracı değişkenler. Psikolojiden aşina olduğunuz bir başka yapı, "susuzluk" kavramıdır. Örneğin, bir araştırmacı katılımcıların bir sıvıdan yoksun kaldıkları saatleri değişimleyebilir ve belirli bir süre sonra içilen sıvı miktarını ölçebilir. Yoksunluk zamanı ile sıvı alınmasına izin verilmesi arasındaki zamanda katılımcıların "susuz" olduklarını söyleyebiliriz –vücut sıvısını tamamlama ihtiyacına karşılık gelen psikolojik deneyim. Susuzluk teorisyenlere, sıvıdan yoksun kalınan süre (bağımsız değişken) ile tüketilen sıvı miktarı (bağımlı değişken) gibi değişkenleri birbirine bağlama olanağı veren bir yapı, bir kavramdır. Susuzluk gibi aracı değişkenler, sadece bağımsız ve bağımlı değişkenleri birbirine bağlamakla kalmazlar, fakat aynı zamanda, değişkenlerin *neden* ilişkili olduklarını açıklamak için de kullanılırlar. Bu yüzden, araştırmacılar bulgularını açıklamak için teorileri kullandıkları zaman, aracı değişkenler önemli bir rol oynar.

Aracı değişkenler ve teoriler yararlıdır, çünkü araştırmacıların görünüşte farklı değişkenler arasındaki ilişkileri belirlemelerine olanak sağlar. Muhtemelen, diğer bağımsız değişkenler "susuzluk"u etkiler (bk. Şekil 2.6). Örneğin, farklı bir bağımsız değişkeni ele alın: Tüketilen tuz miktarı. Yüzeysel olarak, bu iki bağımsız değişken –sıvıdan yoksun kalma süresi ile tüketilen tuz miktarı– çok farklıdır. Böyle olmakla birlikte, her ikisi de daha sonra tüketilen sıvı miktarını etkiler ve susuzluk aracı değişkeniyle açıklanabilir. Diğer bağımsız değişkenler, alıştırma miktarı ile sıcaklık dâhil, sıvı tüketimi ile ilişkilidir; daha fazla egzersiz ve daha yüksek sıcaklık, insanların daha fazla "susuzluk" çekmelerine ve daha fazla sıvı tüketmelerine neden olur. Bu örnekler bağımsız değişkenleri vurgulamakla birlikte, bağımlı değişkenlerin de teori geliştirmede önemli bir rol oynadıklarını belirtmek gerekir. Bu nedenle, bağımlı değişken olarak "sıvı tüketimi"ni ölçmekten ziyade, yaratıcı araştırmacılar susuzluğun psikolojik deneyimi ile ilişkili diğer etkileri ölçebilirler. Örneğin, sıvıdan yoksun kalındığı zaman, bireyler sıvıya erişmek için daha fazla çaba sarf edebilirler veya tadı acı olan suyu

ŞEKİL 2.6 Potansiyel bağımsız değişkenler (solda), potansiyel bağımlı değişkenler (sağda) olarak "susuzluk" aracı değişkenini etkileyebilir.



içebilirler. Bu nedenle, sıvı elde etme çabası veya sıvıdaki acılığın miktarı da bağımlı değişken olarak ölçülebilir.

Aracı değişkenler, psikolojide teori geliştirmede önemlidir. Bizim örneğimizde, Şekil 2.6'da görülen ve görünüşe göre benzerlikleri olmayan değişkenler, "susuzluk" aracı değişkenine dayanan bir teori altında birleştirilebilir. Psikolojide, aracı değişkenlerin –ve teorilerin– başka birçok örneği vardır. Örneğin "depresyon" aracı değişkeni, depresyona neden olduğu varsayılan faktörler (örneğin, nörolojik faktörler, travmaya maruz kalma gibi) ile çeşitli semptomları (örneğin, üzüntü, ümitsizlik, uyku ve iştahsızlık gibi) birbirine bağlar. Aynı şekilde, bir aracı değişken olarak "bellek", çalışma için harcanan zaman miktarı (veya niteliği) ile bu çalışmadan sonraki bir testte ortaya konulan performans arasındaki ilişkiyi açıklamak için kullanılır. Psikoloji bilginiz ilerledikçe aracı değişkenler, değişkenler arasındaki karmaşık ilişkileri anlaşılır hâle getirecek bir anahtar işlevi görecektir.

Teorileri nasıl değerlendireceğimiz ve test edeceğimiz meselesi, psikolojideki ve felsefedeki en zor meselelerden birisidir (örneğin, Meehl, 1978, 1990a, 1990b; Popper, 1959). Kimble (1989) basit ve açık bir yaklaşım önerdi: "En iyi teori, mantıkî ve ampirik testlerin ateşine karşı ayakta kalabilen teoridir" (s. 498). Bilim insanları bir teoriyi değerlendirmek için ilkin mantıkî olup olmadığını anlamak isterler. Yani, bir teorinin anlamlı olup olmadığını ve önermelerinin çelişkili olup olmadığını değerlendirirler. Teorilerin mantıkî tutarlılığı, bilim camiasının eleştirel gözüyle test edilir.

Kimble'ın (1989) bir teoriyi test etmek için önerdiği ikinci "ateş", bir teoriden elde edilen hipotezleri görgül olarak test etmektir. Bir teoriden elde edilen bir hipotezin başarıyla test edilmesi, o teorinin kabul edilmesi olasılığını arttırır. Ama bir hipotezi test etmenin ve bunun sonucunda bilimsel teorileri desteklemenin veya desteklememenin önünde ciddi engeller vardır. Örneğin, bir teori, özellikle de karmaşık bir teori test edilebilir nitelikte çok sayıda hipotez sağlayabilir. Bir teori, bir tek hipotezin test edilmesiyle yanlışlanamaz. Üstelik, teoriler yeterince tanımlanamayan veya aracı değişkenler ile davranış arasında karmaşık ilişkiler öne süren kavramlar içerebilirler. Böyle teoriler uzun ömürlü olabilirler, ama bunların bilim açısından değeri tartışmalıdır (Meehl, 1978). Bilim camiası en sonunda, bir teorinin herhangi bir testinin kesin olup olmadığını belirler.

Genellikle, *dakik kestirimler* sağlayan teorilerin yararlı olması olasılığı çok daha fazladır (Meehl, 1990a). Örneğin, çocukların tipik şekilde soyut akıl yürütme yeteneğini 12 yaşına kadar kazanacağı kestirimini yapan bir teori, bu yeteneğin 12-20 yaşları arasında ortaya çıkacağı kestirimini yapan bir teoriden daha dakik (ve test edilebilir) bir kestirim yapmış olmaktadır. Bir teori inşa edilip değerlendirilirken, bilim insanları ekonomiye büyük önem verirler (Marx, 1963). *Ekonomi kuralı* demek, mümkün açıklamaların en basitini kabul etmek demektir. Bilim insanları, olayları en basit şekilde açıklayan teorileri tercih ederler.

Özetle, iyi bir bilimsel teori, en titiz testleri geçebilen teoridir. Biraz sezgilerimize aykırı olmakla birlikte, titiz test işlemleri, araştırmacılar bir teorinin önermelerini *yanlışlamak* (falsify) istedikleri zaman, doğrulamak istedikleri durumdan daha fazla bilgi sağlar (Shadish, Cook ve Campbell, 2002). Her ne kadar özel bir teorinin önermelerini doğrulayan (confirmation) testler o teoriyi destekler ise de doğrulama, mantikî olarak aynı olayın alternatif açıklamalarını

(teorileri) ortadan kaldırmaz. Yanlışlama testleri bir teorinin kurumuş dallarını budamanın en iyi yoludur. Bilimsel teorileri inşa etme ve değerlendirme bilimsel etkinliklerin merkezidir ve psikoloji biliminin sağlıklı şekilde gelişmesi için esastır.

ÖZET

Bir bilgi elde etme yaklaşımı olarak bilimsel yöntemin temel özelliği, sezgiden ziyade görgül işlemler kullanması ve bir olaydan sorumlu olduğu düşünülen faktörleri kontrol etmeye çalışmasıdır. Bilim insanları en üst düzeyde kontrole, deneysel yöntemle ulaşırlar. Bir deneyde, davranış üzerindeki etkilerini belirlemek için sistematik şekilde değişimleşen faktörlere bağımsız değişkenler denir. Bağımsız değişkenlerin etkisini (eğer varsa) değerlendirmek için kullanılan davranış ölçümlerine ise bağımlı değişkenler denir.

Bilim insanları bulgularını yansız ve nesnel bir şekilde bildirmek isterler. Kavramları işevuruk şekilde tanımlamak, onların bu amaca daha fazla yaklaşmalarına yardım eder. Psikoloji alanındaki araştırmacılar, kavram karşılığı olarak "yapı" kullanırlar. Bilim insanları ayrıca, mümkün olduğunca doğru ve dakik olmak için araçlar da kullanırlar. Olaylar hem fiziksel hem de psikolojik ölçüler kullanılarak ölçülür. Bilim insanları hem geçerli hem de güvenilir ölçümler yapmak isterler. Olayların geçici açıklamasına hipotez denir. Ancak hipotezlerin bilim insanlarına yararlı olabilmeleri için test edilebilir olmaları gerekir. Açık bir tanımı olmayan, kendisini kendisiyle tanımlayan veya bilimin alanı dışındaki düşüncelere veya güçlere başvuran hipotezler, test edilebilir değildir. Hipotezler çoğunlukla teorilerden elde edilir.

Bilimsel yöntemin amaçları betimleme, kestirim yapma, açıklama ve uygulamadır. Hem nitel hem de nicel araştırmalar, davranışı betimlemek için kullanılır. Gözlem, bilimsel betimlemenin temel prensibidir. İki ölçüm korelasyonlu ise, bir ölçümün değerini bildiğimiz zaman diğerinin değerini kestirebiliriz. Bir olayın nedenleri keşfedildiği zaman, o olay anlaşılır ve açıklanır. Bunun için olayların birlikte değiştiğine, bir zaman-sıra ilişkisi olduğuna ve alternatif açıklamaların bertaraf edildiğine dair kanıtların elde edilmesi gerekir. Potansiyel olarak etkili iki değişken birlikte değiştiği zaman, yani her bir değişkenin davranış üzerindeki bağımsız etkisi belirlenemediği zaman, o araştırmada karıştırıcı etki olduğunu söyleriz. Dikkatle düzenlenmiş bir deney araştırmacıya nedensel bir çıkarım yapma olanağı verse bile, o araştırmanın bulgularının başka insanlara ve ortamlara ne ölçüde genelleştirileceği meselesi çözülmüş sayılmaz. Uygulamalı araştırmalarda psikologlar, elde ettikleri bilgileri ve kullandıkları yöntemleri insanların hayatını iyileştirmek için kullanırlar. Öte yandan temel araştırmalar, davranış ve zihinsel süreçler hakkında bilgi elde etmek ve teorileri test etmek için yapılır.

Bilimsel teoriler kurma ve bunları test etme çalışmaları, psikolojiye bilimsel yaklaşımın merkezinde yer alır. Bir teori olayları tanımlayan, bu olaylar arasındaki ilişkileri betimleyen ve olayların nasıl meydana geldiklerini açıklayan mantıkî tutarlılığı olan önermeler setidir. Teorilerin, görgül bulguları organize etme ve test edilebilir hipotezler sağlayarak araştırmalara yol açma gibi önemli

işlevleri vardır. Aracı değişkenler psikolojide teori geliştirmede önemlidir çünkü bu yapılar araştırmacılara bağımsız değişkenler ile bağımlı değişkenler arasındaki ilişkileri açıklama olanağı sağlar.

TEMEL KAVRAMLAR

kontrol 30
bağımlı değişken 31
deney 31
bağımsız değişken 32
yapı 33
işevuruk tanımlama 34
geçerlik 38
güvenirlik 38
korelasyon 45
nedensel çıkarım 47
karıştırma 48
temel araştırma 49
uygulamalı araştırma 49
teori 50

GÖZDEN GEÇİRME SORULARI

1 Şu özelliklerin her birisi için, bilgiye bilimsel yaklaşım ile günlük yaklaşım arasındaki farkı belirtin: Genel yaklaşım ve tutum, gözlem, kavramlar, bildirim, araçlar ve hipotezler.
2 Bir bağımsız değişken ile bağımlı değişken arasındaki farkları belirtin ve her birisi için bir deneyde kullanılabilecek bir örnek verin.
3 İşevuruk tanım kullanmanın psikolojideki avantajları nelerdir? İşevuruk tanımın iki kusurunu açıklayın.
4 Bir ölçme aracının doğruluğu ile dakikliği arasındaki farkı açıklayın.
5 Bir ölçümün geçerliği ile güvenirliği arasındaki farkı açıklayın.
6 Hangi üç tip hipotez, en önemli özellik olan test edilebilirlikten yoksundur?
7 Bilimsel yöntemin dört amacını ve her birisi ile neyin gerçekleştirilmek istenildiğini belirtin.
8 Araştırılan kişi ve yapılmak istenilen genelleştirmenin niteliği açısından, nomotetik ve idiografik yaklaşımlar arasındaki farkları belirtin.
9 Nitel ve nicel araştırmalar arasındaki iki farkı belirtin.
10 İki değişkenin korelasyonlu olduğunu bildikleri zaman, araştırmacılar ne yapabilirler?
11 Kitabınızdan, nedensel bir ilişkinin kurulabilmesi için gerekli olan üç koşulun her birisini gösteren bir araştırma örneği bulun.
12 Temel araştırma ile uygulamalı araştırma arasındaki fark nedir?
13 Aracı değişken ne demektir? "Aşağılama" (var/yok) ile "saldırgan tepkiler" arasında bir aracı değişken olabilecek bir yapı (kavram) önerin. Sizin aracı değişkeninizi içeren bir hipotez kurarak, bu değişkenlerin nasıl ilişkili olabileceklerini açıklayın.
14 Bir bilimsel teoriyi değerlendirmede mantıkî tutarlılığın ve görgül olarak test etmenin rol nü betimleyin.
15 Bir teorinin önermelerini titiz bir şekilde yanlışlamaya çalışmanın, o teoriyi doğrulamaya çalışmaya kıyasla neden daha fazla bilgi verebileceğini açıklayın.

DÜŞÜNDÜRÜCÜ SORULAR

1 Aşağıda betimlenen araştırmaların her birisinde bağımsız değişken(ler)i belirlemeniz gerekiyor. Ayrıca, her bir araştırmada en azından bir bağımlı değişken de belirleyeceksiniz.

A Bir psikolog, gıda yoksunluğunun motor hareketler üzerindeki etkilerini belirlemek istiyor. Bu amaçla 60 sıçanı, dört koşula atıyor. Bu dört koşul, sıçanların gıdadan yoksun kaldıkları zamanın uzunluğu açısından farklıdır: Sıfır saat, 8 saat, 16 saat, ve 24 saat. Araştırmacı daha sonra, farelerin kafesteki faaliyet tekerleğinde harcadıkları zamanı ölçtü.

B Bir beden eğitimi öğretmeni, çocuklar çok sayıda donanımın (örneğin, kaydırak, salıncak, tırmanma duvarı) olduğu geniş bir oyun alanında deneyim kazanırken meydana gelen motor koordinasyonu belirlemek istiyor. Okul öncesi çocukların 8 hafta boyunca haftada 4, 6, veya 8 saat bu ekipmanlarla oynamalarına izin veriliyor. Öğretmen daha sonra, öğrencilerin atlamalarını, zıplamalarını ve tek ayak üzerinde durmalarını isteyerek onların motor koordinasyonlarını test ediyor.

C Bir gelişim psikologu 3 yaş çocuklarında, yanlarında kimin olduğuna bağlı olarak ortaya çıkan sözel davranış miktarını belirlemek istiyor. Bu çocuklar bir laboratuvar ortamında 30 dakika boyunca gözlendiler. Araştırmanın bir koşulunda çocukların yarısının yanında bir yetişkin vardı. Diğer koşulda ise incelenecek çocukların yarısının yanında bir başka küçük çocuk vardı. Psikolog incelenmekte olan her bir çocuğun sözel ifadelerinin sayısını, süresini, ve karmaşıklığını ölçtü.

2 Bir psikolog, "gruplarıyla özdeşleşen bireylerin, tanımadıkları kimselere yardım etme olasılığının daha düşük olduğu" hipotezini test etmek için bir deney yaptı. Üniversite öğrencilerinin, öğrenci etkinlik merkezine yakın bir yerde, "kampüsteki deneyimleri ile ilişkili kısa bir anketi cevaplamaları" istendi. İlk test işlemi yarıyılın başında yapıldı. Kendi üniversiteleriyle özdeşleşme eğilimlerini harekete geçirmek için kendilerine bir kâğıt verildi ve bu üniversitede en çok sevdikleri üç şeyi yazmaları istendi (yüksek özdeşleşme düzeyi). Yirmi öğrenci test edildi. İkinci test işlemi, dönemin son haftasında aynı yerde iki defa öğleden sonraları yapıldı. Bu kontrol koşulunda (özdeşleşme düzeyi düşük), 20 yeni öğrenciden teneffüste yapmayı planladıkları üç şeyi yazmaları istendi. Koşulların her ikisinde de, her bir katılımcı yazdığı kağıdı psikologa verdikten hemen sonra, üzerinde rakip üniversitenin adı yazılı olan bir kazak giymiş bir öğrenci asistan yanlarından geçti. Ama bu anlaşmalı öğrenci, kazara (güya) elindeki dosyayı yere düşürdü ve kâğıtlar ortalığa saçıldı. Beklendiği gibi sonuçlar gösterdi ki yüksek özdeşleşme koşulundaki öğrenciler, dökülen kâğıtları toplaması için anlaşmalı öğrenciye kontrol koşulundan daha az yardımcı oldular. Araştırmacılar, bir grupla özdeşleşmenin, yabancılara (grup dışında olanlara) yardım olasılığını azalttığı sonucuna vardılar.

A Psikoloğun etkisini anlamak istediği bağımsız değişkeni (ve düzeyini) ve bağımlı değişkeni bulun.

B Potansiyel olarak, psikoloğun bağımsız değişkeni ile karışmış olabilecek başka hangi değişken(ler) olabilir? Karışmanın nasıl meydana geldiğini açıklayın ve özdeşleşmenin yardım davranışına etkisi hakkında varılabilecek sonuçları betimleyin.

C Bu deney başka nasıl yapılabilirdi? Öyle ki deneyci, grupla özdeşleşmenin etkileri ve bir yabancıya yardım hakkında açık bir sonuca varabilsin.

3 Mart 2013'te yayımlanmış olan bir habere göre, bir grup araştırmacı tüm dünyada obeziteden kaynaklanan 180.000 ölüm vakasına ulaştı (ABD'deki 25.000 vaka dâhil). Ölümler kola ve enerji içecekleri gibi şekerli meşrubatlarla bağlantılıydı. Dünya Sağlık Teşkilatının (WHO) Global Burden of Diseases'in 2010 verilerinden aldığı bilgileri kullanan araştırmacılar şeker hastalığına, kalp ve damar hastalıklarına, ve kansere bağlı obezitenin neden olduğu ölümleri araştırdılar. Araştırmacılar ayrıca, bu sağlık araştırmasına dâhil edilen ülkelerde kişi başına düşen şekerli meşrubat miktarına da ulaştılar. Sonuçta, içilen şekerli meşrubat miktarı arttıkça, obezite kaynaklı ölüm tehlikesinin de arttığı bulundu. Araştırmacılara göre her 100 obez ölümünden birisinin nedeni, şekerli meşrubatlardır. Önde gelen beslenme uzmanları şekerli meşrubatın, ABD'deki obezite salgınının başta gelen nedeni olduğu sonucuna vardılar. Bazı hükümetler bu araştırmanın verilerine dayanarak, meşrubat tüketimini sınırlama çağrısı yaptılar (Örneğin, New York Bloomberg yasası).

A Araştırmacılara göre şekerli içecek tüketimi, obezite bağlantılı ölüm tehlikesini arttırmakta ve bu tip ölümleri azaltmanın ilk adımı, bu içeceklerin tüketimini azaltmaktır. Bu araştırmanın hangi bulguları, bu nedensel çıkarımı yapabilmek için yeterlidir ve hangi kanıtlar yoktur?

B Bu bulgular hakkında bir sonuca varmadan önce bu özetin dışında başka hangi kaynakları kontrol etmek istersiniz? (Şu site ile başlayabilirsiniz: www.cnn.com/2013/03/19/health/sugary-drinks-deaths.)

4 Bir eğitim psikolojisi dersinde bir öğretim yöntemi olarak "bireysel tepki aygıtı" (clicker) kullanmanın, test başarısını arttırıp arttırmadığını belirlemek için bir araştırma yapılmıştır (Mayer ve ark., 2009). Tepki aygıtı kullanılan sınıfta öğrenciler (2005–2006 öğretim yılı) ders esnasında çoktan seçmeli soruları cevaplamak için bu aygıtı kullandılar. Kağıt-kalem kullanılan sınıfta ise öğrenciler (2006–2007 öğretim yılı) çoktan seçmeli soruları cevaplamak için bir kâğıt-kalem formatı kullandılar. Kontrol koşulunda öğretmen, derste çoktan seçmeli soru sormadı. Ara sınav ve final sınavı sonuçları gösterdi ki bireysel tepki aygıtı kullanan sınıf hem kâğıt kalem formatı kullanan sınıftan hem de kontrol sınıfından daha başarılı oldu. Araştırmacılar, derste tıklama aygıtı kullanmanın öğrencilerin sınav başarısını arttırdığı sonucuna vardılar ve tıklama aygıtının, öğrencilerin ders esnasında uygun bilişsel işleme (processing) yapabilmelerine yardımcı olduğunu öne sürdüler.

A Bu araştırmada, öğretim yöntemi ile öğrencilerin test başarısı arasında nedensel bir çıkarım yapmayı mümkün kılan hangi kanıtlar vardır, hangi kanıtlar yoktur? (Nedensel bir çıkarım yapabilmek için gerekli üç koşulu göz önüne almayı unutmayın.)

B Bilimsel yöntemin dört amacını düşünün ve bu araştırmadaki bulguların bu dört amacı karşılayıp karşılamadığını açıklayın.

İleri Alıştırmaların Cevapları

1 Bu araştırmadaki bağımsız değişken, katılımcıların el-göz koordinasyon görevini tamamladıktan sonra yaşadıkları duygu koşuludur. Üç düzey vardı: minnet, olumlu duygu ve nötr durum. Bağımlı değişken, katılımcıların araştırmacının işbirlikçisi (confederate) olan öğrenciye yardım etme süresiydi (dakika olarak).

2 Araştırma bulgularının alternatif açıklaması şudur: İşbirlikçi bilgisayar problemini çözdüğü zaman, katılımcılar kendilerini daha iyi hissettiler ve bu nedenle, deneyin sonunda daha fazla yardım ettiler. Minnet duygusunun önemli olduğunu göstermek amacıyla araştırmacılar, olumlu duygu durumunu kontrol etmek için bir deneysel koşul (eğlenceli video koşulu) kullandılar. Yani, eğer daha fazla yardımın nedeni sadece olumlu duygu olsaydı, bu katılımcıların da daha fazla yardım etmesi gerekirdi. Sadece minnet koşulundaki katılımcılar en fazla yardımı yaptıkları için, araştırmacılar, özellikle minnetin daha fazla yardıma neden olduğunu öne sürebildiler.

Düşündürücü Soruların Cevapları

A Bağımsız değişken: Dört farklı düzeydeki gıda yoksunluğu süresi (0, 8, 16, 24); bağımlı değişken: Hayvanların faaliyet tekerleğinde harcadıkları zaman (dakika olarak).

B Bağımsız değişken: Haftada üç değişik zaman diliminde (4, 6, 8 saat) oyun alanındaki araçlarda harcanan zaman; bağımlı değişken: Motor koordinasyon testinde alınan puanlar.

C Bağımsız değişken: İki düzeydeki fazladan insan (yetişkin, çocuk); bağımlı değişken: Çocuğun sözel ifadesinin sayısı, süresi, ve karmaşıklığı.

BÖLÜM ÜÇ

Psikoloji Araştırmalarında Etik Meseleler

BÖLÜMÜN ANA HATLARI

GİRİŞ
ARAŞTIRMAYA BAŞLAMADAN ÖNCE GÖZ ÖNÜNE ALINMASI GEREKEN ETİK MESELELER
RİSK/FAYDA ORANI
Risk Belirleme
En Az (Minimal) Risk
Riskin Üstesinden Gelme
BİLGİLENDİRİLMİŞ ONAM
PSİKOLOJİ ARAŞTIRMALARINDA ALDATMA
ARAŞTIRMA SONRASI BİLGİLENDİRME (DEBRIFING)
HAYVANLARLA ARAŞTIRMA
PSİKOLOJİ ARAŞTIRMALARININ RAPOR EDİLMESİ
ETİK AÇIDAN UYGUN BİR ARAŞTIRMA YAPMANIN AŞAMALARI
ÖZET

GİRİŞ

İyi araştırmayı ancak iyi bilim insanları yapabilir. Bilim insanının meslekî yeterliliği ve dürüstlüğü, nitelikli bir bilimin güvencesidir. Bireysel olarak bilim insanı ve bilim camiası (örneğin, APA ve APS gibi profesyonel organizasyonlar), bilimsel çalışmaların doğru bir şekilde yapılması için sorumluluğu paylaşırlar. Bireysel olarak her bilim insanı, bilgiye ulaşmaya ve insanların hayatını iyileştirmeye çalışırken etik sorumlulukla da karşı karşıyadır. Bu genel prensip, bilim insanına çeşitli sorumluluklar yükler. Bilim insanları

- araştırmalarını yetkin bir şekilde yapmalıdır;
- araştırma sonuçlarını doğru şekilde bildirmelidir;
- araştırma kaynaklarını dürüst şekilde kullanmalıdır;
- kendi araştırmalarına ve düşüncelerine katkıda bulunan meslektaşlarını açıkça belirtmelidir;
- araştırmanın toplum açısından sonuçlarını göz önüne almalıdır;
- bilgi ve uzmanlığı ile ilişkili toplumsal endişelerini açıkça söylemelidir (Diener ve Crandall, 1978).

Bilim insanı bu yükümlülüklerini yerine getirmeye çalışırken, bazen zorlu ve belirsiz etik meseleler ve sorularla karşılaşır. Psikologların bireysel olarak etik kararlar verebilmelerine yardımcı olmak amacıyla, APA bir etik yönetmeliği oluşturdu. Bu yönetmelik araştırma veya tedavi yapan veya öğretici veya yönetici olarak görev yapan psikologlar için etik davranış standartları koymuştur (bk. American Psychological Association, 2002; 2010a). Bu etik yönetmeliği cinsel taciz, psikolojik danışmanlık ücretleri, medyada toplumu aydınlatma, test hazırlama ve sınıfta öğreticilik gibi çok çeşitli konularda standartlar içermektedir. Bu yönetmeliğin belirttiği davranış standartlarına ve ideallerine uygun davranmaya çalışmak, tüm psikoloji öğrencileri için de önemlidir. APA'nın sitesine girerek Etik Yönetmelik hakkında bilgi edinebilirsiniz: www.apa.org/ethics.

APA'nın Etik Yönetmeliği'ndeki standartların çoğu, doğrudan hem insan hem de hayvanlar üzerinde yapılan psikoloji araştırmalarını ele alır (özellikle bk. yönetmeliğin 8.01 ile 8.15 arasındaki standartları). Etik yönetmeliklerin çoğunda olduğu gibi, bu standartlar genel bir şekilde ifade edilmiştir; bu nedenle, özel bağlamlarda yeniden yorumlanmaları gerekir. Özel bir araştırma durumunda birden fazla etik standart uygulanabilir ve bazen standartlar birbirleriyle çelişiyor gibi görünebilir. Örneğin, etik bir araştırma insan katılımcıların fiziksel yaralanmalara karşı korunmasını gerektirir. Ama uyuşturucu veya diğer zararlı işlemleri içeren bir araştırma, katılımcıları fiziksel olarak tehlikeye atabilir. Hayvan katılımcıların sağlığı korunmalıdır, ama belirli tipteki araştırmalar acı veren işlemler içerebilir. Bu etik ikilemleri (dilemma) çözmek her zaman kolay değildir ve etik çözüm kararları verilirken incelikli, merhametli bir yaklaşım gerektirir.

İnternet, birçok araştırmacının araştırma şeklini değiştirdi ve bu durum psikologlar için de geçerlidir. Örneğin, dünyanın her yerinden araştırmacılar çoğunlukla bilimsel projeler üzerinde işbirliği yapmaktadırlar ve düşüncelerini ve bulgularını internet aracılığıyla kolayca ve çabucak birbirlerine iletebilmektedirler. Binlerce internet sitesinde, çok geniş bir araştırma arşivi bulunmaktadır.

Araştırmacılar WWW (World Wide Web) aracılığıyla dünyanın her yerindeki insan katılımcılardan veri toplayabildikleri için, bir tek araştırmaya potansiyel olarak *milyonlarca* insanın katılması mümkündür. İnternette psikoloji alanında yapılan araştırma tiplerinden bazıları basit gözlem (örneğin, bir sohbet odasındaki "davranışların" kaydedilmesi), anket (soru listeleri, kişilik testleri), ve değişkenleri değişimlemeyi içeren deneylerdir (bk. Skitka ve Sargis, 2005).

İnternet, davranış alanında çalışan araştırmacılara büyük fırsatlar vermekle birlikte, aynı zamanda birçok etik sorun da doğurmaktadır. Bunların en önemlileri şunlardır: Araştırmacının araştırma ortamında bulunmaması, bilgilendirilmiş onam alınmasında ve araştırma sonrası bilgilendirmede yetersizlik, ve katılımcının gizliliğini korumadaki endişelerdir (bu sorunlara ilişkin incelemeler ve bazı çözüm önerileri için özellikle bk., Buchanan ve Williams, 2010; Kraut ve ark., 2004 ve Nosek, Banaji ve Greenwald, 2002). Biz bu etik meselelerin bazılarını bu bölümde tartışacağız ve sonraki bölümlerde, internette yapılan araştırmaları kullandığımız zaman, tartışmaya devam edeceğiz.

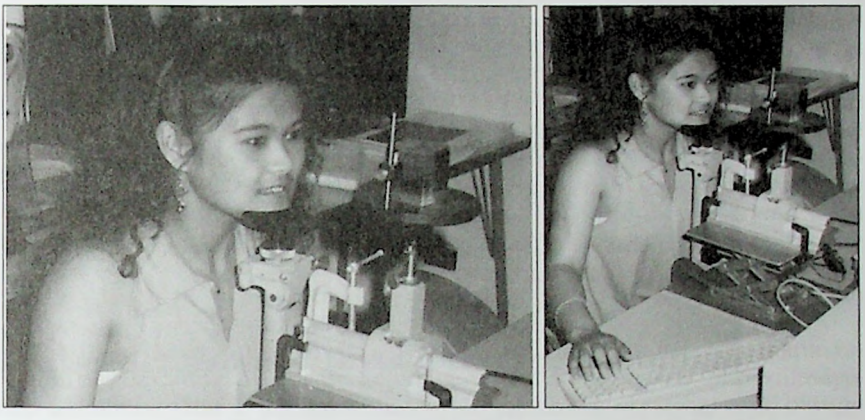
Etik konularda en iyi kararlar, başkalarıyla yapılan görüşmelerden sonra verilir. Bu "başkaları"na arkadaşlar, ama özellikle de belirli bir alanda daha deneyimli ve bilgili olduğu kabul edilen kimseler dâhildir. Gerçekte bir araştırma planını, o araştırmaya katılmamış olan insanlarla incelemek, çoğu durumda yasal olarak gereklidir. Bu bölümün kalan kısmında, Etik Yönetmeliğin özellikle psikoloji araştırmalarını ele alan standartları üzerinde yorumlar yapacağız. Bu tartışmadan sonra, etik sorulara yol açan çeşitli kurgusal araştırma senaryolarını ele alcağız. Kendinizi, bu araştırma önerilerinde ortaya çıkan etik meseleler hakkında karar vermek zorunda olan birisinin yerine koyarak, belirli etik standartları uygulamada ortaya çıkan zorlukların üstesinden gelmenin ve genel olarak etik kararlar vermenin zorluklarını öğreneceksiniz. Bu araştırma önerilerini meslektaşlarınızla, öğreticilerle, ve psikoloji alanında araştırma deneyimi olan başkalarıyla tartışmanızı öneriyoruz.

ARAŞTIRMAYA BAŞLAMADAN ÖNCE GÖZ ÖNÜNE ALINMASI GEREKEN ETİK MESELELER

- Herhangi bir araştırmayı yapmadan önce, o araştırmanın etik standartları karşılayıp karşılamadığı belirlenmeli, salt bu amaçla incelenmelidir.
- Kurumsal İnceleme Kurulu (Institutional Review Boards - IRBs), insan katılımcıların haklarını ve sağlığını korumak amacıyla, psikoloji araştırmalarını incelemektedir.
- Kurumsal Hayvan Bakım ve Kullanım Kurulu (Institutional Animal Care and Use Committees - IACUCs), hayvanlara insanca davranılmasını sağlamak amacıyla, hayvanlar üzerinde yapılan araştırmaları incelemektedir.

Araştırmacılar etik meseleleri, bir araştırma projesine başlamadan önce ele almalıdırlar. Eğer bir araştırma *yapılmaya başlanmadan önce* dikkatle planlanırsa, uygun insanlara ve gruplara danışılırsa etik sorunlar önlenebilir. Bir araştırma etik prensiplere uygun şekilde yapılamazsa, tüm bilimsel süreci zayıflatabilir, bilginin ilerlemesini engelleyebilir ve toplumun bilim insanlarına ve kuruluşlarına olan saygısını azaltabilir. Bu aynı zamanda, bireysel olarak araştırmacıların ve bilim kuruluşlarının ceza almalarına neden olabilir. Psikoloji araştırmalarına başlarken araştırmacıların atmaları gereken önemli bir adım, araştırmaları için kurumsal onay almaktır.

ŞEKİL 3.1 Araştırma insanlar üzerinde yapıldığı zaman, birçok etik sorun ortaya çıkar.

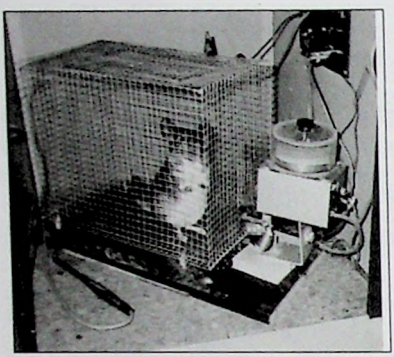


1974 teki Ulusal Araştırma Yasası, biyomedikal ve davranışsal araştırmalarda insan katılımcıların korunabilmesi için bir ulusal komisyonun kurulması ile sonuçlanmıştır. Bu yasa araştırma kurumlarının, başka kuruluşlarca parasal olarak desteklenmesi istenilen araştırmaları incelemeleri için komiteler kurmalarını gerektirmektedir. Üniversiteler bu komiteleri kurdular: *Kurumsal İnceleme Kurulları* (Institutional Review Boards -IRBs). IRBs için yapılan federal düzenlemeyi, şu sitede inceleyebilirsiniz: www.hhs.gov/ohrp. Bir kurumun IRB incelemesi şu anlama gelir: Araştırmacıların, katılımcılara zarar vermemeleri ve onların haklarını korumaları güvence altına alınmıştır (bk. Şekil 3.1). Federal düzenlemeler, üyelerinden çok spesifik taleplerde bulunur ve kurula spesifik görevler yükler (bk. *Federal Register*, June 23, 2005). Örneğin, bir IRB farklı geçmişi ve uzmanlık alanı olan en az beş üyeden oluşmalıdır. Bu komitelerde hem bilim insanları hem de bilim insanı olmayanlar yer almalıdır, ve kurumla ilişkisi olmayan en az bir üye bulunmalıdır. Örneğin, toplum hizmeti yapan din adamı, avukat ve hemşireler, sıklıkla bu komitelerde görev yapar.

IRB, bir araştırmayı onaylama, onaylamama veya değişiklikler yapılmasını isteme yetkisine sahiptir. IRB ayrıca kurumun bakış açısını, araştırmacıyı, ve katılımcıları göz önüne alarak araştırma önerilerini uygun bir şekilde inceleme gibi bir etik sorumluluk altındadır (Chastain ve Landrum, 1999).

ABD Tarım Bakanlığı ile Halk Sağlığı Servisi 1985 yılında laboratuvar hayvanlarının bakımı için yeni bir yönerge hazırladılar (bk. Şekil 3.2). Sonuç olarak şimdi, hayvanlarla araştırma yapan kurumların, hayvanların nasıl bakılacağını ve kullanılacağını belirleyen bir kurullarının yani, *Kurumsal Hayvan Bakımı ve Kullanım Kurullarının* (Institutional Animal Care and Use Committees - IACUC) olması gerekmektedir. Bu komitelerde en az bir bilim insanı, bir veteriner, ve bir de kuruma bağlı olmayan kişinin olması gerekmektedir. IACUC'un hayvan araştırmaları incelemesi, sadece araştırma projesini gözetlemesinin çok ötesine geçer. Hayvan araştırmalarının yapılmasını düzenleyen federal yasalar,

ŞEKİL 3.2 Araştırmalarda hayvan kullanmaya ilişkin etik yönerge, hayvanlara testten önce, test esnasında ve testten sonra nasıl davranılması gerektiğini açıklar.





hayvanların yaşama yerlerinin belirlenmesine ve o hayvanlarla çalışacak olan personelin yeterli eğitimi almasına kadar uzanır (Holden, 1987).

Hemen hemen her üniversite, kurumda yapılacak olan tüm araştırmaların, veri toplamaya başlamadan önce bağımsız bir komite tarafından incelenmesini istemektedir. İnsan veya hayvan katılımcıların yer aldığı araştırmaların incelenmesine ilişkin federal düzenlemelerin ihlali, bir kurumdaki tüm araştırmaların durdurulmasına, federal fonların kesilmesine ve önemli cezaların verilmesine neden olabilir (Holden, 1987). *Araştırma yapmak isteyen herhangi bir kimse araştırmasına başlamadan önce, doğru kurumsal inceleme işlemleri hakkında yetkililerden bilgi almalıdır.* Bir IRB'ye (McCallum, 2001; Newman, 2008; Pollic, 2007) veya bir IACUC'a (LeBlanc, 2001), bir araştırma önerisi sunulduğu zaman, yararlı öneriler alınması olanağı vardır.

RİSK/FAYDA ORANI

- Bir araştırma projesinin öznel bir risk/fayda değerlendirmesi, o araştırmanın yapılıp yapılmaması gerektiğine karar vermek için kullanılır.

Gerekli etik prensiplere uyulup uyulmadığını kontrol etmenin dışında bir IRB, o araştırmanın *risk/fayda oranını* da inceler. Yeni bilgiler elde edildiği zaman ve insanların hayatını iyileştirecek tedaviler bulunduğu zaman, toplum ve bireyler araştırmalardan yarar sağlar. Ama araştırma *yapılmadığı* zaman, bunun bir bedeli de vardır: Bilgi elde etme ve insanların koşullarını değiştirme fırsatını kaçırmış oluruz. Araştırma esnasında eğer katılımcılar zarar görürlerse, bunun onlar için de maliyeti yüksek olur. Elbette araştırmanın başındaki kişi, bu potansiyel maliyet ve fayda hesabını yapan ilk kişi olmalıdır. Bir IRB, araştırmayla herhangi bir çıkar ilişkisi olmayan bilgili kimselerden oluşmalıdır. Böylece bir IRB, risk/fayda oranını belirlemek ve en sonunda da önerilen araştırmayı onaylayıp onaylamamaya karar vermek için daha uygun bir konumdadır.

Temel Kavram

Risk/fayda oranı, "Bu araştırmayı yapmaya değer mi?" sorusunu sorar. Risk/fayda oranının matematik bir cevabı yoktur. Bunun yerine IRB üyeleri

hem birey olarak katılımcılar hem de toplum için risk ve fayda oranını *öznel* olarak değerlendirirler ve *"Araştırmanın faydası, riskinden fazla mıdır?"* sorusunu sorarlar. Risk, potansiyel faydadan daha ağır bastığı zaman, IRB araştırmayı onaylamaz; ama araştırmanın sağlayacağı fayda, riskinden daha ağır bastığı zaman, araştırmayı onaylar.

Bir araştırma etkinliğinde gerçek bir risk/fayda dengesiyle ilgili verilecek kararı birçok faktör etkiler. En temel faktör, araştırmanın potansiyel bilimsel ve toplumsal değerinin yanı sıra, katılımcılar için taşıdığı risk ve potansiyel olarak onlara sağlayacağı faydadır (Fisher ve Fryberg, 1994). Eğer bireylere açık ve yakın bir faydası öngörülüyorsa veya araştırmanın açık bir bilimsel ve sosyal değeri varsa, daha büyük bir risk göze alınabilir. Örneğin, psikotik davranış için yeni bir tedavi araştıran bir araştırma projesi, katılımcıların risk almalarını gerektirebilir. Ama eğer önerilen tedavinin fayda sağlama olasılığı yüksek ise, hem bireyler hem de toplum için sağlayacağı fayda, araştırmada göze alınan riskten daha ağır basabilir.

Risk/fayda oranını belirlemede araştırmacılar, aynı zamanda araştırmanın niteliğini de göz önüne alırlar; yani, geçerli ve anlamlı sonuçlar elde dilip edilmeyeceği göz önüne alınır. Daha açık bir ifadeyle, "Eğer bilimsel niteliği zayıf olduğu için bir araştırmadan iyi bir sonuç çıkmayacaksa, o araştırmada katılımcının zamanını, dikkatini, emeğini ve ayrıca parayı, araç gereci harcamayı nasıl savunabiliriz?" (Rosenthal, 1994b, s. 128). Bu nedenle *bir araştırmacı etik olarak, bilimsel standartları en üst düzeyde karşılayacak nitelikte araştırmalar yapmakla yükümlüdür.*

Potansiyel bir risk olduğu zaman bir araştırmacı, daha az riskli başka bir seçenek olmadığı için o yöntemi kullandığına dair komiteyi ikna etmelidir. Araştırmacı aynı zamanda, önceki araştırmacıların o araştırma konusunu başarılı bir şekilde ele almadıklarından da emin olmalıdır. Bir araştırmacı, psikoloji literatürünü iyice incelemezse, önceden yapılmış olan ve bu yüzden de katılımcıları gereksiz yere riske atan bir araştırma yapabilir.

Risk Belirleme

- Psikoloji araştırmalarındaki potansiyel riskler fiziksel yaralanma, sosyal ilişkilere zarar ve zihinsel veya duygusal strestir.
- Risk değerlendirmesi yapılırken potansiyel katılımcıların günlük etkinlikleri, onların zihinsel ve fiziksel sağlıkları ve yetenekleri göz önüne alınmalıdır.

Katılımcıların "risk altında" olup olmadıklarını belirleme, etik kararlar vermede işe karışan zorluklar gösterir. Hayatın, kendisi risklidir. İşe veya okula gidiş geliş, caddeden karşıya geçiş, ve asansörlere binme hep risk unsuru taşır. Basitçe, gösteri amaçlı bir psikoloji deneyi yapmak bile bazı riskler taşır. Psikoloji deneylerinde katılımcıların asla risk altında olmamaları gerektiğini savunmak, araştırmaları tamamen engeller. Neyin risk teşkil edeceğine dair kararlar, günlük hayatın bir parçası olan riskleri göz önüne almalıdır.

Araştırmacılar risk belirlerken aynı zamanda, katılımcıların özelliklerini de göz önüne almalıdırlar. Belirli etkinlikler bazı insanlar için ciddi bir risk teşkil edebilir ama başka bazıları için etmez. Merdivenlerde koşmak yaşlı insanlar için kalp krizi olasılığını arttırabilir ama muhtemelen genç yetişkinlerin çoğu için böyle bir risk yoktur. Aynı şekilde, çok depresif ve endişeli kimseler belirli psikolojik görevlere, böyle olmayan kimselerden çok daha şiddetli tepkide bulunabilirler.

Bu nedenle, risk hesaplanırken araştırmacıların araştırmaya katılma olasılığı olan özel grupları veya bireyleri hesaba katmaları gerekir.

Biz riski çoğunlukla fiziksel yaralanma olasılığı açısından düşünürüz. Ama sıklıkla, sosyal bilim araştırmalarına katılan katılımcılar sosyal veya psikolojik zarar görebilirler. Örneğin, eğer araştırmacının dışında olan insanlar bir katılımcının kişisel bilgilerini öğrenirlerse, katılımcı mahcup olabilir. Psikoloji araştırmalarında zekâ, kişilik özellikleri, politik, sosyal ve dinsel inanç, ve özel davranışlar hakkında bilgi toplanabilir. Bir katılımcı muhtemelen bu gibi bilgilerinin öğretmenleri, çalışanlar veya arkadaşları tarafından bilinmesini istemez. Katılımcıların verdikleri cevapların gizliliğini korumada başarısızlık, sosyal zarar olasılığını arttırabilir.

İnternette yapılan araştırmalarda katılımcılar için muhtemel bir risk, katılımcıya ait olduğu anlaşılabilecek bilgilerin, araştırma ortamı dışındaki insanlar tarafından öğrenilmesi olasılığıdır (Kraut ve ark., 2004). İnternetin gizliliği sağlama gücü olmakla birlikte (Nosek ve ark., s. 165), araştırmacılar verileri letmede, saklamada ve katılımcılarla araştırma sonrası etkiletişimde gizliliği korumak için önlem almalıdırlar (ayrıca bk. Buchanan ve Williams, 2010; Pollick, 2007).

Bazı psikoloji araştırmaları, eğer katılımcılar ciddi zihinsel ve duygusal stres yaşarlarsa, psikolojik risk taşıyabilirler. Bir katılımcının beklemekte olduğu odaya sigara dumanı girdiği zaman yaşadığı stresi düşünün. Araştırmacı, bir acil durum varmış izlenimi yaratmak için sigara içebilir. Ama sigara dumanının bu gerçek nedeni ortaya çıkıncaya kadar, katılımcı sıkıntı yaşayabilir. Duygusal veya psikolojik stresin ne zaman meydana geleceğini önceden kestirmek her zaman kolay değildir.

Araştırmacılar çocuk istismarı ve kişiler arası şiddet konusunda bilgi toplamaya çalıştıkları zaman ortaya çıkan ikilemi ele alalım (bakınız Becker-Blease ve Freyd, 2006). Katılımcılardan geçmişte yaşadıkları çocuk istismarına veya aile şiddetine örnekler vermelerini isteme, duygusal olarak stresli olabilir. Yine de çoğu araştırmacı, bu gibi deneyimlere ait bilgilerin, bilim insanının hem toplumun boşanma, okul başarısızlığı, suçluluk gibi bazı sorunlarına hem de klinik araştırmalara dair anlayışını geliştirmesine yardım ettiği düşüncesindedir. Fakat nasıl ve ne zaman? Becker–Blease ve Freyd uğranılan istismar hakkında soru sormanın ve sormamanın etik yönlerini tartıştılar. Bu araştırmacılar, soru sormamanın bazı sakıncalarına dikkat çektiler. Örneğin, bilimin bu konuda ilerlemesi engellenmiş olabilir; mağdurların yardım almaları engellenebilir, istismara karşı nasıl davranacaklarını öğrenmelerini engelleyebilir ve nihayet, kendilerine yardımcı olabilecek toplumsal güçlerden habersiz kalabilirler. Çocuk istismarı araştırmaları aynı zamanda, istismar konusundaki tabuyu yıkmaya ve kurbanları, bu tartışmaların önemli olabileceğini anlamaya yöneltebilir. Becker-Blease ve Freyd'e göre soru sormaktan kaçınma "istismarcılara yardım eder, kurbanları incitir." (s. 225). Bu nedenle, soru sormamanın maliyeti, bir risk/fayda analizi açısından tartılmalıdır.

Bir psikoloji deneyine katılma, bazı insanlar için kendi başına bir endişe kaynağıdır. Bir öğrenci katılımcı bir defasında *HUP, LEK* gibi anlamsız hecelerden oluşan bir listeyi öğrendikten sonra, araştırmacının kendisi hakkında çok şey öğrendiğinden emin olduğunu söyledi. Bu öğrenciye göre araştırmacı, kendisinin listeyi öğrenirken kullandığı kelime çağrışımlarını inceleyerek, onun kişiliği

hakkında bilgi edinmeyi amaçladı. Gerçekte ise bu öğrenci, unutmayı ölçmek için düzenlenmiş basit bir bellek deneyine katılmıştır. *Bir araştırmacı, katılımcılarını duygusal ve zihinsel stresten korumak zorundadır; mümkün olduğunda, katılımcının deneydeki görevi hakkında edindiği yanlış anlamadan kaynaklanan stres de buna dâhildir.*

En Az (Minimal) Risk

- Araştırmadaki işlemler veya etkinlikler katılımcının her günkü hayatında yaşadıklarına benzer olduğu zaman, bir araştırmanın "en az risk" içerdiği düşünülür.

Temel Kavram

Bir ayrım bazen, "risk altındaki" bir katılımcı ile "en az risk altında" olan bir katılımcı arasında yapılır. **Minimal risk** demek, katılımcının araştırma esnasında karşılaşacağı zarar veya yaşayacağı rahatsızlığın, günlük hayatında veya rutin fiziksel veya psikolojik testlerde yaşayabileceği zarar veya rahatsızlıktan *daha fazla olmaması* demektir. Bu testlerin bir örneği, birçok psikoloji laboratuvarı araştırmasında, çeşitli zihinsel yetenekleri değerlendirmek için kullanılan çok uzun testlerin bilgisayarla uygulanmasıdır. Katılımcıların bu testleri hızlı bir şekilde tamamlamaları istenebilir ve performansları hakkında geri bildirim verilir. Bu durumda stres altında kalmaları olasılığı olmakla birlikte, yaşayacakları psikolojik zarar tehlikesi, muhtemelen sınavda yaşadıklarından daha fazla değildir. Bu nedenle, bu gibi araştırmalar üniversite öğrencileri için ancak minimal risk taşır. Zarar olasılığının minimal'den büyük olduğuna karar verildiği zaman, bireylerin *risk altında* olduğu düşünülür; bu durumda araştırmacı, katılımcıların sağlığını korumak için daha fazla önlem almakla yükümlüdür.

Riskin Üstesinden Gelme

- Katılımcı ister "risk altında" isterse "minimal risk altında" olsun, korunması gerekir. Risk arttıkça, daha fazla yardım gerekir.
- Katılımcıları sosyal ilişkilerinde yaşayabilecekleri riskten korumak için, sağladıkları verilere isimleri yazılmamalıdır veya, eğer bu mümkün değilse, verdikleri bilgilerin gizliliği korunmalıdır.

Potansiyel risk küçük olsa bile, yine de araştırmacının riski en aza indirmeye çalışması ve katılımcıları koruması gerekir. Örneğin, bir bellek testinden önce basitçe bu testin zekâyı veya kişiliği ölçmediğini belirtmek bile, bazı katılımcıların yaşayabilecekleri stresi azaltabilir. Muhtemel zararın günlük hayatta karşılaşılandan önemli ölçüde fazla olduğuna karar verilen durumlarda, araştırmacının katılımcıyı koruma yükümlülüğü de artar. Örneğin, bir psikoloji deneyinde katılımcının ciddi bir duygusal stres yaşaması olasılığı olduğu zaman, *Kurumsal İnceleme Kurulu*'na göre (Institutional Review Board – IRB) bir danışmanın katılımcının yanında olması ve kendisine araştırmadaki deneyimleri hakkında yardımcı olması gerekir. Sizin de kolayca hayal edebileceğiniz gibi, internette yapılan araştırmalar bu açıdan zor etik ikilemler doğurur. Katılımcılar tıpkı laboratuvar araştırmalarında olduğu gibi, internet ortamında da duygusal stres yaşayabilirler. Ama araştırma ortamında olmadıkları için, yaşadıkları stresi ve zararı yönetmekte ve azaltmakta yeterince başarılı olamazlar. Bu soruna bir çözüm yaklaşımı şu olabilir: Risk altında olan katılımcıları belirlemek amacıyla ön veriler elde edilebilir ve bu katılımcılar asıl araştırmanın

İLERİ ALIŞTIRMA I

Aşağıdaki araştırma durumlarının her birisinde, katılımcıların "minimal (yani, günlük hayattakinden fazla olmayan) risk" altında mı yoksa "risk" altında mı olduğuna karar verin. Eğer katılımcıların "risk altında" olduklarına karar verirseniz, bu riski azaltmak için araştırmacıya hangi önerileri yapabileceğinizi düşünün. Bunu yaparken şüphesiz, bu bölümde tartışılacak olan etik meselelerin bazılarına girmiş olacaksınız

1 Üniversite öğrencilerine, kendilerinin o anki duygu durumlarını betimleyen bir sıfat tarama testi veriliyor. Araştırmacının amacı, depresyonda olan öğrencileri belirlemektir; bu durumda olanları, daha sonra depresyonun neden olduğu bilişsel sorunları inceleyen bir araştırmaya almak istiyor.

2 Bir bakımevindeki yaşlılara, etkinlik odasında bir başarı testi veriliyor. Araştırmacının amacı, ilerleyen yaşla birlikte zihinsel bir gerileme olup olmadığını belirlemektir.

3 Psikolojide araştırma yöntemleri dersindeki öğrenciler, dersin ortasında bir başka öğrencinin sınıfa girdiğini görüyorlar; bu öğrenci öğretmenle bağırarak ve öfkeyle konuşuyor ve sonra sınıftan çıkıp gidiyor. Bu olay, görgü tanıklığı dersinin bir parçası olarak, dersin öğretmeni tarafından önceden planlanmıştır; olaydan hemen sonra öğretmen, sınıftaki öğrencilerden, sınıfa girip çıkan kişinin kıyafetini betimlemelerini istiyor.

4 Bir araştırmacı, psikolojiye giriş dersi öğrencilerinin bir araştırmaya katılmalarını istiyor; söz konusu araştırmada, alkolün bilişsel işlevler üzerindeki etkisi incelenecektir. Bu araştırmada bazı öğrencilerin, bir bilgisayar oyunu oynamaya başlamadan önce, meyve suyuyla karıştırılmış olarak yaklaşık 60 gram alkol almaları gerekiyor.

dışında tutulabilir. Bununla birlikte, yüksek risk içeren araştırmalar, etik olarak, internette yapılmayabilir (Kraut ve ark., 2004).

Katılımcılar için minimal düzeyden daha fazla risk taşıyan bir araştırma, ancak aynı verileri daha düşük riskle elde edebilecek yöntemler yoksa, yapılabilir. Bazı durumlarda deneysel işlemler yerine, gözlem veya anket kullanan betimsel yaklaşımlar kullanılmalıdır. Araştırmacılar aynı zamanda, deneysel stres taşımayan ve "doğal olarak meydana gelen" işlemlerden de faydalanabilirler. Birçok insan çok önemli yakınlarını kaybeder ve buna benzer travmatik olaylar yaşar. Psikoloji araştırmalarının önemli ilgi alanlarından birisi de insanların bu gibi ağır travmaların üstesinden nasıl geldikleri meselesidir (Bonanno, 2004). Bu araştırmaların özel örnekleri işten çıkarılmalar ve şirket küçülmeleri (Armstrong-Stassen, 1994; 2006), kasırganın yol açtığı hasar (Anderson, 1970) ve terorist saldırılardır (Back, Kufner, ve Egloff, 2010; Skitka, Bauman ve Mullen, 2004; Solomon, Gelkopf ve Bleich, 2005).

Katılımcıları sosyal çevrelerinde uğrayabilecekleri zarardan korumak için araştırmacılar, katılımcıların bilgilerini gizli tutmalıdır; bu amaçla isimlerini ve kim olduklarını belli edecek başka bilgileri istememelidir. Bu mümkün olmadığı zaman araştırmacı, kaydettiği kimlik bilgilerini araştırma esnasında silerek, verilerin gizliliğini korumalıdır. Araştırmacının katılımcıları bir defadan fazla test etmesi gereken durumlar olabilir ki bu durumda katılımcıları daha sonra tekrar bulması gerekir. Veya alınan bilgiler hassas bilgiler olabilir. Bu gibi durumlarda, araştırmanın başında, hangi verilerin hangi katılımcılara ait olduğu kodlu olarak kaydedilmeli ve katılımcılara seçkisiz (random) olarak tahsis edilmelidir ki hangi bilgilerin kime ait olduğunu araştırmacıdan başkası öğrenemesin. Katılımcıların cevap kağıtlarının üzerinde sadece bu kodlu numaralar bulunmalıdır. Hangi kodun hangi katılımcıya ait olduğunu kaydettiği ana liste kilit altında korunmalı ve araştırmacıdan başkası bu listeye erişememelidir. İnternet aracılığıyla araştırma yapanlar, gizli dinleme ve verilere korsan ulaşma olasılıklarına karşı özellikle dikkatli olmalı ve sosyal riskleri en aza indirmek için uygun tüm önlemleri almalıdırlar (bk. Kraut ve ark., 2004).

Katılımcıların, verdikleri cevapların isimsiz ve gizli tutulduğundan emin olmaları, araştırmacı için de çok yararlıdır, çünkü bu durumda katılımcı araştırmacıya daha çok güvenir ve daha doğru cevaplar verir (Blanck, Bellack, Rosnow, Rotheram-Borus, ve Schooler, 1992). Verdiği bilgilere araştırmacıdan başka kimsenin ulaşamayacağından emin olduğu zaman, katılımcının yanlış bilgi vermesi veya bilgi saklaması olasılığı azalır.

BİLGİLENDİRİLMİŞ ONAM

- Araştırmacılar ve katılımcılar, çoğunlukla bir bilgilendirilmiş onay işlemi kullanarak aralarında bir sosyal anlaşma yaparlar.
- Araştırmacı etik olarak, araştırmada uygulanacak işlemleri açık şekilde betimlemekle yükümlüdür; araştırmanın, katılımcının katılımını etkileyebilecek herhangi bir özelliği varsa bunu belirtmek ve katılımcının araştırmaya dair herhangi bir sorusunu cevaplamak zorundadır.
- Katılımcıya, herhangi bir ceza almaksızın, istediği zaman araştırmadan çekilme hakkı tanınmalıdır.
- Araştırmaya katılmaları için kimseye baskı yapılmamalıdır.
- Katılımcılar etik olarak, araştırma esnasında doğru bilgi vererek, aldatmayarak veya başka hilelere başvurmayarak uygun şekilde davranmak zorundadır.
- Onay verme yeterliliği olmayan küçük çocuklar veya zihinsel engelli bireyler gibi katılımcılardan da onay alınmalıdır; bu gibi katılımcıların onayları, onların yasal koruyucularından alınmalıdır.
- Araştırmacılar, bilgilendirilmiş onamın gerekip gerekmediğine karar verecekleri zaman, örneğin araştırma halka açık ortamlarda yapılacağı zaman, bir IRB dâhil, bilgi sahibi başka kimselerle de görüşülmelidir. Bu gibi ortamlar, katılımcıların gizliliğini korumak için özel bir dikkat gerektirir.
- Gizlilik, katılımcıların kendileri hakkındaki bilgilerin başkalarına nasıl bildirileceğine karar verme hakkı demektir.

Temel Kavram

Etik Yönetmeliğinin araştırmayla ilgili önemli bir kısmı, bilgilendirilmiş onamla ilgili meselelere ayrılmıştır. Bu yerinde bir karardır, çünkü bilgilendirilmiş onam araştırmacı ile katılımcı arasındaki sözleşmenin önemli bir parçasıdır. **Bilgilendirilmiş onam**, bir kimsenin bir araştırmanın niteliğini, katılmadığı zaman ne ile karşılaşacağını ve kendisinin katılma isteğini etkileyebilecek tüm faktörleri anladıktan sonra açık şekilde ifade ettiği katılma arzusudur (bk. Şekil 3.3).

Araştırmacı, katılımcının araştırmayla ilgili herhangi bir sorusunu cevaplamak için yeterince çaba göstermeli ve araştırma esnasında onun haklarına ve itibarına saygılı olmalıdır. Böylece katılımcılar, katılımları hakkında onam vermeye karar verebilirler. Katılımcılar onamlarını, uygunsuz teşvikler veya baskılar olmaksızın, özgürce vermelidir. Katılımcılar aynı zamanda, herhangi bir ceza veya olumsuz kanaat ile karşılaşmaksızın, onamlarını geri çekebileceklerini de bilmelidirler. *Yazılı bilgilendirilmiş onamı, katılımcılar en az riskten fazlasına maruz kalacakları zaman, kesinlikle gereklidir.*

ŞEKİL 3.3 Halk Sağlığı Dairesi, 1932-1972 arasında Alabama eyaletinin Macon kentinde yaşayan ve tedavi görmeyen yoksul zenci erkeklerde frengi hastalığının gelişimini araştırdı. Bu araştırma yapılırken, kendilerinden bilgilendirilmiş onam alınmamış olan bu erkeklerin frengiye yakalandıklarından bile haberleri yoktu ve tedavi edilmeden bırakılmışlardı. Hayatta kalanlarını Clinton yönetimi fark etti.



Bir araştırmaya katılmayı kabul eden katılımcılar, aynı zamanda, etik açıdan uygun şekilde davranmakla yükümlüdürler. Örneğin, katılımcılar yönergeye dikkat etmeli ve araştırmacının söylediği şekilde davranmalıdırlar. Taylor ve Shepperd (1996), katılımcılar kendilerine söylenilen şekilde davranmadıkları zaman neler olacağını gösteren bir araştırma bildirdiler. Bu araştırmada araştırmacı, kendi aralarında deney hakkında konuşmamaları için uyardıktan sonra, katılımcıları kısa bir süreliğine yalnız bıraktı. Ama katılımcılar yalnız kaldıktan sonra deney hakkında konuşmaya ve birbirleriyle bilgi paylaşmaya başladılar; tabi bu durum deneyin değerini ortadan kaldırdı. Üstelik, deneyci sonradan katılımcılara araştırmanın işlemleri ve amaçları hakkında ne bildiklerini sorduğunda, usulsüz konuşmaları esnasında araştırma hakkında önemli bilgiler elde ettiklerinden söz etmediler. Bu örnek, daha geniş bir prensibi gösteriyor: *Katılımcıların yalan söyleme, aldatma, veya diğer hileli davranışları yapmaları, araştırmaların bilimsel değerini düşürür.*

Bazı kimselerden gerçek bir bilgilendirilmiş onam alınamaz; bunlar zihinsel engelli veya duygu durum bozukluğu olanlar, küçük çocuklar, ve araştırmanın niteliğini ve muhtemel riskleri kavramada zorluk çekenlerdir (bk. Şekil, 3.4). Bu

ŞEKİL 3.4 Bilgilendirilmiş onam meselesi, özellikle çocuk katılımcılarla çalışılacağı zaman önemlidir.



durumlarda bilgilendirilmiş onam katılımcıların ebeveynlerinden veya yasal koruyucularından alınmalıdır. Ama mümkün olan her durumda "rıza", yani katılıma gönüllü olma isteği, katılımcıların kendisinden alınmalıdır.

İnternette yapılan araştırmalarda, bilgilendirilmiş onam almak ayrı etik sorunlara yol açar. Şu durumu ele alalım: Çoğu durumda internetten katılan araştırmacı yönergeyi okuduğunu, anladığını ve araştırmaya katılmaya gönüllü olduğunu gösteren seçeneği işaretler; ama bu işaretleme, katılımcıyı yasal olarak bağlayan bir "imza" yerine geçer mi? Bir araştırmacı, bu şekilde onay verenlerin gerçekten de yönergede belirtilen yaşta olup olmadıklarını veya bilgilendirilmiş onam yazısını tam olarak anlayıp anlamadıklarını nasıl bilebilir? Katılımcının bilgilendirilmiş onam açıklamasını anlayıp anlamadığını belirlemenin bir yolu olarak şu önerilebilir: Açıklamanın içeriği hakkında kısa bir test uygulamak; çocukları yetişkinlerden ayırt etmek için ise sadece yetişkinlerin bilebilecekleri nitelikte bilgiler dâhil etmek (Kraut ve ark., 2004). Her ne zaman böyle etik ikilemler ortaya çıksa, deneyimli ve bilgili insanlardan yardım istemek akıllıca bir yöntem olur, ama *etik bir araştırma yürütmenin nihaî sorumluluğu yine araştırmacınındır*.

Katılımcı bulmak için başvurulan yöntemlerden hangilerinin usulsüz bir teşvik veya baskı oluşturduğuna karar vermek, her zaman kolay değildir. Bir saatlik bir psikoloji deneyine katılmaları için üniversite öğrencilerine 9 dolar ödemek, genellikle uygunsuz bir baskı kabul edilmez. Sokaklardaki çok yoksul, evsiz insanları saati 9 dolara bir araştırmaya dâhil etmek, daha baskıcı ve daha az kabul edilebilir bir yöntem olabilir (Kelman, 1972). Mahkûmlar bir

KUTU 3.1

BİLGİLENDİRİLMİŞ ONAM ALMAYA DAİR ESASLAR

Uygun bir bilgilendirilmiş onam araştırmanın amacını, araştırmacının kim olduğunu ve kimlerle birlikte çalıştığını ve araştırma boyunca hangi işlemlerin yapılacağını açıkça belirtmelidir. Katılımcılar onay formunu okuduktan ve sorularının cevaplarını aldıktan sonra, onay formuna tarih atılmalı ve araştırmacı ve katılımcı tarafından imzalanmalıdır. Araştırmalarda İnsan Katılımcıları Koruma Bürosu (The federal Office for Human Research Protections - OHRP) araştırmacıların bilgilendirilmiş onam hazırlamalarına yardımcı olacak esasları yayımlamıştır. Biz bu prensipleri uyarlarken, onay formunda bulunması gereken daha başka hususları da formumuza dâhil ederiz. Bir IRB (institutional review board) ek bilgiler isteyebilir. OHRP prensiplerinin tam metnine ve bu konuyla ilgili önemli federal dokümanlara şu bağlantıdan erişilebilir: www.hhs.gov/ohrp/humansubjects/guidance/ictips.htm.

- Araştırma alanınıza özgü ve teknik terimler kullanmaktan kaçının; bilgilendirilmiş onam metni, katılımcıların kolayca anlayabilecekleri bir dille, açık şekilde yazılmalıdır.
- Birinci tekil şahıs (örneğin, "Ben biliyorum ki ...", veya "Bence ..." gibi) kullanmaktan kaçının, çünkü böyle bir ifade öneri olarak anlaşılabilir ve yanlış bir şekilde, kesin bir bilgiymiş gibi kullanılabilir. "Eğer katılmayı kabul ediyorsan, şunları, şunları yapman istenecektir" gibi ifadeler tercih edilmelidir. Bu bilgilendirilmiş onam metnini, bir yasal belge gibi değil bir öğretim aracı gibi düşünün.
- Katılımcının araştırmada karşılaşacağı ve yapacağı şeylerin hepsini, hangi nitelikte olduklarını (örneğin, hangi anlamda deneysel), önceden tahmin edilebilecek zararlarını, verebileceği rahatsızlıkları, aksaklıkları, ve riskleri betimleyin.
- Katılımcının bu araştırmaya katılmakla kazandığı şeyleri kendisine açıklayın. Eğer kazancı sadece topluma veya bilime yardım ise, bu açıklanmalıdır.
- Araştırmaya katılımın alternatiflerini açıklayın. Eğer üniversite öğrencilerinden oluşan bir "katılımcı havuzu" kullanılıyorsa, o zaman psikoloji araştırmaları hakkında bilgi edinmenin alternatif yolları açıklanmalıdır.
- Katılımcılara, tanınabilecekleri kişisel bilgilerin ne şekilde gizli tutulacağı açıklanmalıdır. Çok hassas bilgilerin toplandığı durumlarda bir IRB, Gizlilik Sertifikası gibi ek koruma önlemleri isteyebilir.
- Eğer araştırma nedeniyle minimal düzeyden fazla bir zarar olasılığı varsa, gönüllü telafi ve tedavi yapılacağını belirtin.
- Katılımcının yasal hakları korunmalıdır.
- Araştırma hakkında bilgi sahibi olan bir "irtibat görevlisi" belirlenmelidir; öyle ki, araştırma sonrasında soruları olan katılımcıların sorularını cevaplayabilsin. Sorular şu üç alandan birisine ilişkin olabilir ve *bu alanlar onam sürecinde ve dokümanlarında açık şekilde ifade edilmelidir*: Araştırma deneyimi, katılımcıların hakları ve araştırmayla ilgili zararlar. Bazen, örneğin katılımcıyı IRB'ye veya bir kurumsal temsilciye sevk eden birden fazla irtibat görevlisi de olabilir.
- Onam belgesinde bir gönüllü katılım ifadesi yer almalıdır; bu ifadede, katılımcının hiçbir ceza almadan, istediği zaman araştırmadan çekilebileceği belirtilmelidir.

psikoloji deneyine katılmayı reddederlerse, yöneticilerin kendilerinin işbirliğine yatkın olmadıklarını düşüneceklerini, bu nedenle şartlı tahliye olmakta zorluk çıkaracaklarını düşünebilirler. Gerçekten de federal yönetmelik (45 CFR 46.305) araştırmacıların mahkûmlara, katılmamaları hâlinde şartlı tahliyelerinin bundan *etkilenmeyeceğini* açıklamalarını istemektedir.

Üniversite öğrencilerinin dersin bir parçası olarak sınıfta yapılacak psikoloji deneylerinde katılımcı olmaları istendiği zaman (bu deneyim muhtemelen kendileri için bir eğitim değeri de taşımaktadır), bu araştırmaya katılmak istemeyenlere seçenekler sunulmalıdır. Bu alternatif seçeneklerin gerektirdiği zaman ve çaba, araştırmaya katılanlarınkiyle denk olmalıdır. Kullanılan alternatif görevler çoğunlukla araştırma makaleleri özetleme, davranış gözlemleri yapma, mezun öğrencilerin veya fakültenin araştırma sunumlarına katılma ve gönüllü kamu hizmeti yapmadır (bk. Kimmel, 1996).

ŞEKİL 3.5 Hangi davranışın "umumî" hangi davranışın "özel" olduğuna karar vermek her zaman kolay değildir.



IRBs, araştırmacılardan, insan katılımcıların yer aldığı herhangi bir araştırmada uygun bir bilgilendirilmiş onam işlemi yürütüldüğüne dair belge ister. Ama şunu da bilmek önemlidir ki, OHRP'nin ilkelerinin de ifade ettiği gibi, bilgilendirilmiş onam süreci araştırmacı ile gelecekteki katılımcılar arasında aktif bir bilgi paylaşımı süreci olmalıdır (www.hhs/ohrp; bk., "Bilgilendirilmiş onam ile ilgili sık sorulan sorular"). Kutu 3.1'de, "Tüm onam belgelerine uygun" örnek bir formdan ziyade, uygun bir bilgilendirilmiş onam oluşturmanın esaslarını verdik. Uygun onam işlemleri ve yazılı belgeleme duruma ve araştırma evrenine göre biraz değişir. Etik prensiplere uygun ve katılımcıların haklarını koruyan bir bilgilendirilmiş onamın nasıl hazırlanacağı konusunda iyi bir yardım kaynağı IRB üyeleridir.

Bazı durumlarda araştırmacıların bilgilendirilmiş onam almaları gerekmez. Bunun en açık örneği, araştırmacı kamuya açık alanlarda, hiçbir müdahalede bulunmaksızın insanların davranışlarını incelediği durumdur. Örneğin, bir araştırmacı bir üniversite kampüsündeki ırk ilişkileri hakkında bulgu toplamak isteyebilir; bu amaçla, kampüs içinde yürüyen "karışık ırk"tan ve "aynı ırk"tan grupların sıklığını gözleyebilir. Bu gözlemi yapmadan önce bu araştırmacı, öğrencilerden izin almak zorunda değildir. Ama eğer belirli öğrencilerin kimlikleri kaydedilecekse, bilgilendirilmiş onam gerekli olur.

Temel Kavram

Bir davranışın ne zaman umumî, ne zaman özel olduğuna karar vermek her zaman kolay değildir. **Mahremiyet** (Privacy), bireylerin kendileri hakkındaki bilgilerin başkalarıyla nasıl paylaşılacağına karar verme hakkıdır. Hangi bilginin mahrem olduğuna karar verirken araştırmacılara yardımcı olabilecek üç önemli boyut, *bilginin hassasiyeti*, *ortam* ve bilgiyi *yayma şekli*dir (Diener ve Crandall, 1978). Açıktır ki, bazı tür bilgiler, diğerlerinden daha hassastır.

Kendileriyle cinsel hayatları, dinî inançları veya işledikleri suçlar hakkında konuşulan bireylerin, araştırmacının bu bilgileri nasıl kullanacağı hakkında daha fazla endişeye kapılması beklenmelidir.

Bir davranışın özel mi yoksa umumî mi olduğuna karar vermede ortam da etkilidir. Bir konsere katılma gibi bazı davranışlar, kolayca umumî bir davranış olarak nitelenebilir. Umuma açık bir ortamda insanlar belirli bir derecede özel davranmaktan vazgeçerler. Ama bu gibi ortamlarda meydana gelen bazı davranışların özel mi yoksa umumî mi olduğuna karar vermek kolay değildir (bk. Şekil 3.5). Arabanızı kullandığınız zaman, umumi bir tuvaleti kullandığınız zaman veya bir parkta ailece piknik yaparken davranışlarınız özel midir yoksa umumî midir? İnternette sohbet ederken özel bir davranış mı yoksa umumî bir davranış mı sergilemiş oluyorsunuz? Bu gibi durumlardaki verilen etik kararlar, toplanan verilerin niteliğine ve bu bilgilerin nasıl kullanılacağına bağlıdır.

Eğer araştırmada elde edilen bulgular grup ortalamalarıyla veya oranlarla yayımlanacaksa, bireysel bilgilerin açıklanması olasılığı ortadan kalkar. Yok, eğer bireysel bilgiler yayımlanacaksa, katılımcıların gizliliğini korumak için kodlama sistemleri kullanılabilir. *Bireylerin veya grupların hassas bilgilerini, onların izni olmadan, yayımlamak, ciddi bir etik ihlaldir.* Eğer katılımcılar farkında olmadan onlar hakkında hassas bilgiler toplandıysa (örneğin, gizlenmiş bir gözlemci tarafından), araştırmacı gözlemden sonra bu kimselerle temas kurmalı ve bu bilgileri kullanıp kullanamayacağını kendilerine sormalıdır.

İLERİ ALIŞTIRMA II

APA'nın Etik Yönetmeliği'ne göre psikologlar, doğal gözlem araştırmalarında bilgilendirilmiş onam hazırlamayabilirler (bk., Standart 8.05). Ama az önce belirttiğimiz gibi, doğal gözlemin ne zaman "umumî" ortamda yapıldığına karar vermek her zaman kolay değildir. Şu araştırma senaryosunu ele alın ve araştırmacıların araştırma öncesinde bilgilendirilmiş onam almaları gerekip gerekmediğine karar verin. Araştırmacılardan daha fazla bilgi almanız gerekiyor da olabilir. Eğer öyleyse, karar vermeden önce ne tür ek bilgiler istersiniz? Göreceksiniz ki bilgilendirilmiş onam gerekliliği, araştırma ortamını çok etkileyecektir. Örneğin, araştırmacının davranışı "doğal" koşullarda kaydetmesini zorlaştırabilir. Etik karar verme ikilemlerinin bazı örnekleri şunlardır.

1 Üniversite öğrencileri arasında bir içki içme davranışı araştırmasında, bir öğretim üyesi bir öğrenci derneği partisine katılıyor ve öğrencilerin içtikleri içki miktarını kaydediyor.

2 Eşcinsel gruplar üzerinde yapılan bir araştırmanın bir parçası olarak, bir eşcinsel araştırmacı bir eşcinsel beyzbol takımına giriyor; amacı, sezon boyunca takım karşılaşmaları esnasında katılımcıların davranışlarını kaydetmektir. Bütün oyunlar şehrin bir eğlence alanında, herkesin seyrine açık oynanıyor.

3 Kadın ve erkek araştırmacılar tuvalet kabinlerine gizlenerek, kadınların ve erkeklerin umumî tuvaletteki davranışlarını (örneğin, sifonu çekme, el yıkama, çöp kutusuna kâğıt atma, kapılara yazı yazma ve resim çizme gibi) gözlüyor.

4 Üniversite mezunu bir öğrenci, üniversite öğrencilerinin kopya çekme davranışlarını araştırmak istiyor. Çok sayıda öğrencinin katılımıyla büyük bir salonda yapılacak olan bir sınavdan önce, salondaki bir projeksiyon kabininin içine saklanıyor. Görüş açısı sınıfın büyük kısmını kapsıyor ve bir el dürbününün yardımıyla çoğu öğrencinin davranışlarını görebiliyor. Araştırmacı sağa-sola bakmaları, kâğıt değiştirmeleri, not alışverişlerini, cep telefonu kullanımlarını, telefonla mesajlaşmaları ve diğer şüpheli hareketleri kaydediyor.

Araştırmacı, bu durumda izin vermeyi reddeden katılımcıların bilgilerini kullanmamalıdır. Gizlilikle ilgili en zor kararlar, yukarıda bahsedilen üç boyuttan birisinde açık bir etik ihlâlin olduğu ama diğer iki boyutta olmadığı durumlardır veyahut her üçünde de belirsiz bir problemin olduğu durumlardır. Örneğin, bireylerin karartılmış bir sinema salonundaki davranışları, o bireylere ait hassas bilgilerin ortaya çıkması potansiyeli taşır, halbuki böyle bir ortam, pekâlâ umumî bir ortam olarak sınıflanabilir.

Mümkün her durumda, katılımcıların verdikleri bilgilerin ne şekilde gizli tutulacağı kendilerine açıklanmalıdır; öyle ki gizliliklerinin korunması için alınmış olan önlemlerin yeterli olup olmadığına kendileri karar verebilsinler. Bilgilendirilmiş onam hazırlama prensipleri uygulanırken bir denge sağlanmalıdır; onam bir yandan araştırmacıların insan davranışını araştırma ihtiyaçlarını karşılamalı, öte yandan da katılımcıların haklarını korumalıdır.

PSİKOLOJİ ARAŞTIRMALARINDA ALDATMA

- Psikoloji araştırmalarında araştırmacılar, katılımcıdan bilgi sakladıkları zaman veya katılımcıya kasıtlı olarak yanlış bilgi verdikleri zaman aldatma ortaya çıkar. Doğası gereği aldatma, bilgilendirilmiş onamın etik prensiplerinin bir ihlalidir.
- Bazı psikoloji araştırmalarında aldatma, gerekli bir strateji olarak görülmektedir.
- Ama katılımcıları araştırmaya katılmaları için aldatma, daima etik dışı bir davranıştır.
- Araştırmacılar aldatma yöntemine başvurmayı düşündükleri zaman, aldatmanın maliyeti ile araştırmanın kazancı arasındaki farkı dikkatle hesaplamalıdırlar.

Temel Kavram

Araştırmayla ilgili en tartışmalı etik meselelerden birisi, aldatmanın kullanılmasıdır. **Aldatma**, ya katılımcıdan bazı bilgiler *saklanarak* ya da katılımcıya araştırmanın bazı yönleri hakkında kasıtlı şekilde *yanlış bilgi verilerek* yapılabilir. Bazı insanlara göre katılımcılar asla aldatılmamalıdır çünkü etik kurallara göre araştırmacı ile katılımcı arasındaki ilişkinin açık ve dürüst olması gerekir (örneğin, Baumrind, 1985). Bazılarına göre aldatma ahlâk açısından kötüdür; yalan söylemekten farklı değildir. Aldatma, bilgilendirilmiş onam ile çelişmektedir. Ayrıca Baumrand'a göre katılımcıyı aldatmak akıllıca da değildir çünkü katılımcı, araştırmacının "hilekâr" olduğunu ve katılımcılara verdiği bilgilere inanılmaması gerektiğini çabucak anlar. Üstelik aldatma toplum için de zararlıdır, çünkü insanların uzmanlara olan güvenini sarsar ve başarılı birçok çalışmaya karşı onları şüpheciliğe sevk eder.

Ama bütün psikologlar bu görüşte değildir (bk. Christensen, 1988; Kimmel, 1998). Örneğin, Milgram'a (1977) göre aldatma bir çeşit "teknik illüzyon"dur ve bilimin çıkarı için izin verilmelidir. En nihayet gerçek hayatta da insanları bazı şeylere inandırmak için bazen illüzyona başvurulur. İnsanların daha güçlü deneyimler yaşamalarını sağlamak için filmlerde, video oyunlarında, televizyonda görsel ve işitsel etkiler (effects) sıklıkla kullanılır ve insanlar bu etkilerin "gerçek" olmadıklarını bilir. Eğer belirli etkiler gerçek hayatın bir parçası ise, bilim insanı insan davranışını anlamak için neden illüzyonlar kullanamasın?

Milgram, bazı durumlarda ahlâk prensiplerinin askıya alınabileceğine işaret eder. Eğer bir suç işlendiğini öğrendiysek, bunu yetkililere bildirmek zorundayız. Öte yandan, bir zanlı kendi avukatına bilgi vermişse, avukatın bu bilgiyi, müşterisinin suçlu olduğunu ortaya çıkaracak bir bilgi olsa bile, kendisine saklamalıdır. Bir hekim, bedenimiz üzerinde çok özel muayeneler yapar ki bu işlem, doktorun muayenehanesinde meşrudur ama muayenehane dışında kabul edilemez bir davranıştır. Milgram'a göre bilimin çıkarı için, psikologlara bazen ahlâk prensiplerini çiğneme izni verilmelidir.

Son birkaç on yılda araştırmalarda aldatma konusu ön plana çıkmış olmasına rağmen, psikoloji araştırmalarında aldatmanın kullanılması, yaygın bir araştırma tekniği olarak kullanılmaya devam edilmektedir. Örneğin, Skitka ve Sargis (2005), interneti bir veri toplama aracı olarak kullanan sosyal psikologları araştırdı ve bu araştırmaların %27'sinde internet katılımcılarının aldatıldığını belirledi.

Etik açıdan bu kadar tartışmalı olan aldatmanın, yine de bu kadar yaygın kullanılması nasıl açıklanmalıdır? Bunun bir nedeni, araştırmanın bazı yönlerini katılımcıdan saklamadan, bazı araştırmaları yapmanın olanaksız olmasıdır (bk. örneğin, Kimmel, 1998). Başka bazı durumlarda ise, katılımcıların belirli tutum ve davranışları benimsemeleri için kendilerine yanlış bilgi verilmesi gerekmektedir (bk. Kutu 3.2).

KUTU 3.2

ŞOK EDİCİ BİR ALDATMA

Şu araştırmayı ele alın:

Bir "öğrenme deneyi"ne katılmak amacıyla iki kişi bir psikoloji laboratuvarına giriyor. Bir araştırmacı (beyaz önlüklü), bu iki kişiye araştırmanın cezanın öğrenmeye etkisi üzerine olduğunu söylüyor. Bu iki kişi, kimin "öğretmen" kimin "öğrenci" olacağını belirlemek için kâğıt çekiyor. Ama bu iki kişiden birisi, aslında araştırmacıya yardım eden işbirlikçidir (confederate), ve kâğıt çekimi hilelidir, öyle ki gerçek katılımcı hangi kağıdı çekerse çeksin, öğretmen rolü oynayacaktır. Gerçek katılımcı daha sonra "öğrenci"nin yan odaya alındığını, bir sandalyeye kayışlarla bağlandığını, ve bileklerine elektrotlar bağlandığını görür (bakınız, Şekil 3.6). Araştırmacı, öğrencinin kelime listesini öğrenirken her hata yaptığında elektrik şoku alacağını söyler.

Gerçek katılımcı daha sonra laboratuvar odasına alınır; burada, 30 şalter butonu olan etkileyici görünüşte bir şok jeneratörü vardır. Her bir butonun üzerinde kaç voltluk olduğunu gösteren etiketler vardır ve bu değerler 15 volttan 450 volta kadar değişmektedir. Butonların



ŞEKİL 3.6 Milgram'ın 1960'larda yaptığı deneylerde katılımcılara, araştırmaların amacının insanların otoriteye itaatini araştırmak olduğu söylenmedi ve birçok katılımcı, bir diğer insana şiddetli şoklar verilmesini isteyen deneycinin dediklerini yaptı. Bu araştırmanın yeni bir versiyonu için bu bölümdeki Kutu 3.4'e bakın.

üzerinde ayrıca "Hafif Şok", "Güçlü Şok", ve "Dikkat' Şiddetli Şok" ibareleri yazılıdır. Bu son ibareden sonraki iki buton üzerinde sadece XXX işaretleri vardır. Gerçek katılımcıya fikir edinmesi için örnek bir şok verilir. Araştırmacı öğretmene, öğrenci kelime çiftlerini öğrenirken her hata yaptığında şok vermesi ve her hatada şok miktarını arttırması talimatını verir.

Öğrenci, yapacağı hataları ve bu hataların karşılığında verilen şoklara nasıl tepki vereceğini belirten ve önceden hazırlanmış olan yönergeye uygun şekilde davranır. Önce şoktan şikayet eder, sonra protesto eder, en sonunda da şokun verdiği acıyla bağırır. Araştırmacıya, deneye son vermesini söyler, ama deneyci devam eder deneye. Öğretmen 180 voltluk butona geldiği zaman öğrenci, "Artık daha fazla dayanamıyorum" diye feryat eder; 270 voltta ise acıklı bir çığlık atar. Öğrenci 300 voltta "Artık cevap vermeyeceğim" diye bağırır ve çığlık atmaya devam eder. Nihayet 330 voltluk butona basıldığı zaman, öğreniciden hiç ses çıkmaz.

Gerçekte olan şey ise şudur: Öğrenciye hiçbir şok verilmemiştir. Deneydeki asıl bağımlı değişken, öğretmenin (gerçek katılımcının) araştırmacının "emir"lerine uyarak verdiği en yüksek dereceli şoktur. Bu deneydeki işlemin aynısını kullanarak yapılan bir düzineden fazla deneyde, gerçek katılımcıların çoğunluğu araştırmacıya itaat etti ve en yüksek şoku verdi, ama verilen şokun miktarı bazı faktörlere bağlı olarak değişti; örneğin, öğretmenin, öğrencinin sesini duyduğu ama ayrıca kendisini görüp görmediğine göre, veyahut da öğretmenin şok butonuna kendisinin basıp basmadığına göre verilen şokun en üst düzeyi değişti (Milgram, 1974).

Stanley Milgram'ın 1960'larda yaptığı ve günümüzde çok ünlü bu deneylerin amacı, "otoriteye itaat"i incelemektir. Deneylerin rahatsız edici sonuçları gösteriyor ki bir otorite bizden istediği zaman, bir başkasına acı verici şok vermek gibi ahlâken savunulamayacak bir eylemi yapabiliyoruz. Milgram bu deneyleri, insanların itaatinin toplu katliamlara neden olduğu Soykırım'larda (Blass, 2009) neler olduğunu anlamak için yapmıştır.

Milgram'ın araştırmaları, APA'nın Etik Yönetmeliği yürürlüğe girmeden önce ve federal yasalar kurumsal inceleme kurullarının (Institutional Review Board – IRB) insanlar üzerindeki deneyleri izlemelerini zorunlu kılmadan önce yapılmıştır. Bu deneyler, kullanılmış olan aldatmanın ve katılımcıda yarattığı stresin büyüklüğü nedeniyle, etik endişelere yol açmıştır (Elms, 2009). Sonuç olarak, on yıllardan bu yana, Milgram'ın işlemlerini kullanan araştırmalar yapılmadı. Riskler, sağlanan kazançtan daha fazla gibi görünüyor.

Örneğin, Kassin ve Kiechel (1996) insanların gerçekte yapmadıkları bir şeyi yapmışlar gibi itirafta bulunmalarını etkileyen faktörleri araştırdı. Araştırmacıların amacı, zanlıların gerçekte işlemedikleri suçları üstlenmelerine yol açan etkenleri keşfetmekti. Deneyde katılımcıların görevi, sesli olarak okunan harfleri daktilo etmekti. Kendilerine, harfleri daktilo ederken klavyenin Alt tuşuna vurmamaları, çünkü bunun bilgisayarı bozacağı söylendi. Katılımcı yazmaya başladıktan bir süre sonra, araştırmacı bilgisayarı hileli bir şekilde bozdu ve katılımcıyı Alt tuşuna basmakla suçladı. Katılımcıların hiçbirisi Alt tuşuna basmadığı hâlde, %70'i Alt tuşuna bastığına dair yazılı bir itirafı imzaladı. Eğer katılımcılar araştırmanın amacını önceden bilselerdi, muhtemelen imzalamazlardı. Eğer bu deneydeki aldatma bilgilendirilmiş onamda belirtilseydi, insanların yanlış itiraflarda bulunma olasılıklarını araştırmak mümkün olmazdı. APA'nın Etik Yönetmeliği dikkatle okunduğu zaman görülecektir ki aldatmayla ilişkili prensipler (Standart 8.07) "aldatmanın kullanılmasında rekabet eden değerleri ve anlayışları içerir; öyle ki bilimin ilerlemesi, zarardan kaçınma, ve özgür iradeye saygı, hepsi etik dengenin unsurlarıdır" (Behnke, 2009, s. 76; bk. www.apa.org/ethics).

Aldatma, önemli araştırmaların yapılabilmesi için bazen haklı görülebilir ama birden fazla minimal riski olan bir araştırmaya katılmaları için katılımcıları aldatma, daima etik dışıdır. Etik Yönetmelik'te belirtildiği gibi, "*Psikologlar,*

fiziksel acıya veya şiddetli duygusal strese neden olacak bir araştırmaya katılmaları için, katılımcıları aldatamazlar" (Standart 8.07b).

Araştırmanın bir amacı, bireylerin normal davranışlarını araştırmaktır. Aldatmaya başvurmanın temelindeki varsayım şudur: Bir deneyin gerçek niteliğini gizlemek bazen gereklidir; öyle ki katılımcılar her zaman davrandıkları gibi davransınlar veya deneycinin yönergelerine uygun davransınlar. Aldatmanın sık ve gelişigüzel kullanılması, sorunlara yol açabilir. Katılımcılar araştırmacının sık sık aldattığına inanırlarsa, aldatılmayı bekleyebilirler ve araştırma hakkındaki şüpheleri onların normal davranmalarını engelleyebilir (Baumrind, 1985). Aldatmayı savunanlar, katılımcıların aldatılmaya olumsuz bir tepki göstermiyor göründükleri araştırmaları (Burger, 2009; Christensen, 1988; Epley ve Huff, 1998; Kimmel, 1996; Milgram, 1974) ve insanların şüpheciliği artmakla birlikte, toplam etkinin küçük göründüğü araştırmaları işaret ediyorlar (Kimmel, 1998). Aldatmanın sürekli kullanımını savunanlara göre esas mesele şudur ki aldatmanın tüm psikoloji araştırmalarında yasaklanması, araştırmacıların birçok önemli araştırmayı yapmalarını engeller (Kimmel, 1998).

Internet aracılığıyla yapılan araştırmaların sayısı arttıkça araştırmacıların, aldatmanın kullanılmasına özellikle dikkat etmeleri önemlidir; bu sadece toplumun araştırmacılara güvensizlik duymalarına yol açacağı için değil, fakat aynı zamanda aldatmanın bir sistemi (yani interneti) "zehirleme" potansiyeli olduğu için de böyledir; en nihayet insanlar interneti, sosyal destek sağlamak ve başkalarıyla iletişim kurmak için kullanmaktadırlar (Skitka ve Sargis, 2005).

Birçok araştırmacı aldatmanın sürekli kullanılmasından yana olmakla birlikte, şu hususlar çok iyi düşünülmeden aldatma asla kullanılmamalıdır: *(1) Araştırmanın, bilimsel bilgimize katkısı açısından önemi, (2) Aldatma kullanılmasına gerek kalmayacak alternatif yöntemler ve (3) Aldatmanın zararları* (Kelman, 1972). Bu son değerlendirme, aldatmanın derecesini ve katılımcılara zarar verme olasılığını işaret eder. Kelman'a göre "Eğer bir araştırma çok önemliyse ve alternatif yöntemler yoksa, minimum düzeyin üstünde bir aldatmanın kullanılması haklı görülebilir." (s. 997).

ARAŞTIRMA SONRASI BİLGİLENDİRME (DEBRIEFING)

- Araştırmacılar, araştırma bittikten sonra bile, katılımcılara yardımcı olmanın yollarını aramak zorundadır. Bu yolların en iyilerinden birisi, araştırmadan sonra kendilerine araştırma hakkında detaylı bilgi vermektir.
- Araştırma sonrası bilgilendirme, hem katılımcı hem de araştırmacı için yararlıdır.
- Etik açıdan araştırmacılar, neden aldattıklarını uygun ilk fırsatta katılımcılara açıklamak zorundadır.
- Araştırma sonrası açıklama, araştırmanın niteliği ve onların araştırmadaki rolleri hakkında katılımcılara bilgi verir ve katılımcıların araştırma süreci hakkında eğitimine yardımcı olur. Araştırma sonrası bilgilendirmenin esas önemli işlevi, araştırmaya katıldıkları için, katılımcıların kendilerini iyi hissetmelerini sağlamaktır.
- Araştırma sonrası bilgilendirme araştırmacıya, katılımcıların araştırma sürecini nasıl gördüklerini öğrenme, araştırma sonuçlarının içyüzünü kavrama ve ileride yapacağı araştırmalar için düşünceler geliştirme olanağı sağlar.

Yıllar geçtikçe birçok araştırmacı, katılımcıları kendi araştırmalarının ihtiyacı olan verileri toplamak için gerekli bir "nesne" olarak görme tuzağına düştü. Araştırmacılar bazen, son veriyi topladıktan sonra katılımcılara karşı olan sorumluluklarının bittiğini düşündüler. Bir teşekkür veya el sıkışma çoğunlukla, araştırmanın bittiğinin işareti oldu. Oysa katılımcılar muhtemelen, araştırmanın kendisi hakkında ve kendilerinin bu araştırmadaki rolleri hakkında çok belirsiz düşüncelerle ayrıldılar. Araştırma planlanacağı ve yapılacağı zaman, araştırma deneyiminin araştırmadan *sonra* katılımcıları nasıl etkileyeceğini ve katılımcıların bu araştırmadan nasıl faydalanacaklarına dair çareler düşünmek önemlidir. Bu endişeler, APA Etik Yönetmeliği'nin belirtmiş olduğu ahlâk prensiplerinden ikisiyle yakından ilişkilidir: Yardım (kişinin lehine davranma) ve insanların haklarına ve onuruna saygı.

Daha önce dedik ki katılımcıların verdikleri bilgilerin gizliliğini koruma hem katılımcıların lehinedir (sosyal açıdan zarar görmelerini engeller) hem de araştırmacının lehinedir (örneğin, katılımcıların daha dürüst davranmaları olasılığını arttırır). Aynı şekilde, araştırmanın sonunda yapılacak **araştırma sonrası bilgilendirme** de hem katılımcıların hem de araştırmacının lehinedir (Blanck ve ark., 1992). Araştırmada aldatma kullanıldığı zaman, *araştırma sonrası bilgilendirme, neden aldatma yapıldığını katılımcılara açıklamak, araştırmaya katılımları hakkındaki muhtemel yanlış düşüncelerini ele almak ve aldatmanın herhangi bir zararlı etkisini gidermek için gereklidir. Araştırma sonrası bilgilendirmenin aynı zamanda, katılımcıları araştırma konusunda (araştırmanın mantığı, yöntemi, sonuçları gibi) eğitmek ve onların araştırmaya katılımları konusunda olumlu duygularla ayrılmalarını sağlamak gibi önemli işlevleri de vardır.* Araştırmacılar, katılımcıların araştırmaya özel katkıları hakkında daha fazla bilgi edinmelerine fırsat vererek bilimsel sürece daha fazla katılma duygusu yaşamaları için olanak sağlamalıdır. Kassin ve Kiechel'in (1996) yukarıda açıkladığımız yanlış itiraf deneyine katılan katılımcılar, araştırma sonunda bilgilendirme yapıldıktan sonra, araştırmayı anlamlı bulmuş ve araştırmaya kendi katkılarının değerli olduğunu düşünmüşlerdir.

Temel Kavram

Araştırma sonrası bilgilendirme, katılımcıların araştırmaya kendi spesifik katkıları hakkında ve genel olarak araştırma hakkında daha fazla bilgi edinme fırsatı yaratır. Örneğin, katılımcılar kendi performanslarının, araştırmacının yapmalarını istediği şeylerden ve test koşulları gibi faktörlerden nasıl etkilendiğini öğrenebilirler. Psikoloji araştırmalarına katılmanın değerli olduğu düşüncesi, çok sayıdaki psikoloji öğrencisinin araştırmalara katılmalarına gerekçe olarak gösterilmektedir; bu yüzden, üniversite öğrencilerini araştırmaya alan araştırmacılar, bu katılımın bir eğitim değeri taşıdığını göstermek zorundadırlar. Ders öğretmenleri bazen araştırma sonrasındaki bilgilendirmeyi esas alarak öğrencilerinin, araştırmanın amacı, araştırmada kullanılan teknikler ve araştırmanın davranışı anlama açısından önemi hakkındaki düşüncelerini kısa raporlar hâlinde yazmalarını isterler. Böyle raporlar yazan öğrenciler, bunu yapmayanlara kıyasla, kendi araştırma deneyimleri hakkında daha fazla tatmin olmaya ve katılımdan daha kazançlı çıkmaya eğilimlidirler (Richardson, Pegalis ve Britton, 1992).

Araştırma sonrası bilgilendirme araştırmacıların, araştırma süreci hakkında katılımcıların ne düşündüklerini öğrenmelerine yardım eder. Bir araştırmacı, katılımcıların araştırmadaki özel bir işlemi kendi planladığı gibi görüp görmediklerini anlamak isteyebilir (Blanck ve ark., 1992). Örneğin, insanların başarısızlığa nasıl tepkide bulunduklarını öğrenmek isteyen bir araştırma, tamamlanması olanaksız bir görev kullanabilir. Ama eğer katılımcılar kendi performanslarının başarısız olmadığına karar verirlerse, araştırmacının hipotezi test edilemez.

Araştırma sonrası bilgilendirme araştırmacıya, katılımcıların kendi performanslarının başarısız olduğunu düşünüp düşünmediklerini veya o görevi tamamlamanın olanaksız olduğunu kabul edip etmediklerini görme olanağı sağlar.

Katılımcıların araştırmayı nasıl algıladıklarını anlamak istediği zaman araştırmacı, katılımcılara fazla baskı yapmamalıdır. Katılımcılar genellikle bilimsel sürece yardımcı olmak isterler. Katılımcılar, psikoloji araştırmalarında kendilerinden bilgi saklanabileceğini bilebilirler. Hatta araştırma hakkında bildikleri önemli detayları (örneğin, görevi tamamlamanın olanaksız olduğunu) açıklarlarsa, araştırmayı "berbat" edeceklerinden de endişe edebilirler. Bu muhtemel sorunu önlemek için araştırma sonrası bilgilendirmenin sohbet havasında ve dolaylı olması gerekir. Bu çoğunlukla en iyi şekilde, açık uçlu ve genel sorular kullanılarak yapılabilir (örneğin, "Bu araştırmanın hangi konuda yapıldığını düşünüyorsun?" veya "Bu araştırmada yaşadıkların hakkında neler düşündün?" gibi). Araştırmacı daha sonra, araştırmanın işlemleri hakkında spesifik sorular sorarak görüşmeye devam edebilir. Mümkün olduğunca bu spesifik sorular katılımcılara, hangi cevabın beklendiğini belli eden ipuçları vermemelidir (Orne, 1962).

Araştırma sonrası bilgilendirme bir başka açıdan da araştırmacı için yararlıdır: Araştırmanın işlemlerindeki sorunları belirleme ve yapacağı yeni araştırmaları için düşünceler geliştirme fırsatı sağlayabilir (Blanck ve ark., 1992). Araştırma sonrası bilgilendirme aynı zamanda, katılımcıların performansının nedenlerine ilişkin ipuçları da sağlayabilir ki bu ipuçları da araştırma bulgularının yorumlanmasına yardımcı olabilir. Nihayet, katılımcılar bazen araştırma materyalinde hatalar keşfederler –örneğin, eksik bir bilgi veya belirsiz bir ifade; böyle bir şey varsa, araştırma sonrası bilgilendirme sırasında araştırmacıya söylerler. Daha önce de söylediğimiz gibi, *araştırma sonrası bilgilendirme hem katılımcı hem de araştırmacının kendisi için yararlıdır.*

İnternette yapılan araştırmalarda araştırmacının kendisi olmadığı için, uygun bir araştırma sonrası bilgilendirme zor olabilir. Bu konudaki özel sıkıntı şudur ki internet araştırmasına katılan katılımcı istediği zaman kolayca araştırmadan çekilebilir (Kraut ve ark., 2004). Buna bir çözüm önerisi şu olabilir: Araştırma öyle programlanmalıdır ki katılımcı pencereyi zamansız kapattığında otomatik olarak görünsün (Nosek ve ark., 2002). Bir internet araştırmasından sonra araştırmacı, araştırma sonrası bilgilendirme materyalini bir web sitesinde yayımlayabilir ve hatta, yeni bilgiler geldikçe bu materyali güncelleyebilir (bk. Kraut ve ark., 2004). Nihayet bir araştırma tamamlandığı zaman araştırmacı, araştırmanın özet sonuçlarını katılımcılara elektronik postayla gönderebilir; böylece katılımcılar, araştırmanın amaçları ile sonuçlar arasında nasıl bir ilişki olduğunu daha iyi anlayabilirler. Elbette araştırmacılara elektronik postayla bilgi vermek için, onların gizliliğini korumaya yönelik ek önlemler alınması gerekir.

HAYVANLARLA ARAŞTIRMA

- Hayvanlarla araştırma yapılarak, insanlara yararlı (örneğin, hastalıkların tedavisi hakkında) bilgiler elde edilir.
- Etik açıdan araştırmacılar, hayvanların edinilmesi, bakımı, kullanımı, ve araştırmanın sonunda federal yasalara ve mahallî yasalara ve düzenlemelere ve profesyonel standartlara uygun şekilde yok edilmeleri hakkında bilgi edinmek zorundadırlar.
- Araştırmada hayvan kullanımı konusundaki meseleler karmaşıktır ve tartışılmaya devam edilmesi gerekir.

Çok sayıdaki önemli meseleyi aydınlatmak amacıyla, her yıl milyonlarca hayvan laboratuvar araştırmalarında test edilmektedir. Yeni ilaçlar insanlarda kullanılmadan önce, hayvanlar üzerinde denenmektedir. Çevreye yeni girecek maddeler, aynı şekilde, etkilerinin neler olacağını görmek amacıyla önce hayvanlara verilmektedir. Hayvanlara hastalık bulaştırılmaktadır; böylece araştırmacılar hastalığın belirtilerini gözleyebilmekte ve olası tedavileri test edebilmektedirler. Yeni cerrahî işlemler –özellikle de beyin ameliyatları– önce hayvanlar üzerinde denenmektedir. Birçok hayvan, davranış araştırmalarında örneğin etologlar ve deneysel psikologlar tarafından da kullanılmaktadır. Stres ile diyabet arasındaki ilişkinin hayvan modeli, örneğin, diyabette rol oynayan psikosomatik etkenlerin anlaşılmasına yardımcı olmuştur (Surwit ve Williams, 1996). Bu araştırmalar, insanın mutluluğuna katkıda bulunan önemli bilgiler sağlamıştır (Miller, 1985). Ama bu araştırma sürecinde birçok hayvan ağrı, stres, rahatsızlık ve hastalık çekmekte ve ölmektedir. Her ne kadar araştırma laboratuvarlarında en çok kullanılan hayvanlar kemirgenler, özellikle de sıçanlar ve fareler ise de araştırmacılar çok çeşitli hayvanları kullanmaktadırlar; örneğin, maymunlar, balıklar, köpekler, ve kediler bunlardan bazılarıdır. İnsan tepkileri için iyi model oluşturdukları için, bazen özel hayvanlar da laboratuvarlarda kullanılmaktadır. Örneğin işitme ile ilgilenen psikologlar laboratuvarda bazen çinçilla (Amerika tavşanı) kullanırlar, çünkü bu hayvanların işitsel süreçleri insanınkilere benzerdir.

Hayvanları laboratuvarda araştırma denekleri olarak kullanmak, çoğunlukla doğal karşılandı. Gerçekte İncil'in, insanın daha basit tüm canlılar üzerindeki "hakimiyetini" destekleyen ifadeleri, hayvanların laboratuvarda kullanılmalarına destek olarak görülmektedir (Johnson, 1990). Ama hayvanların araştırmalarda kullanılmaları, daha sık şekilde, insanları tehlikeye atmadan daha fazla bilgi edinme ihtiyacına dayandırılmaktadır. Bununla birlikte, laboratuvar araştırmalarında hayvanların rolüne dair birçok soru ortaya çıkmıştır (Novak, 1991; Shapiro, 1998; Ulrich, 1991). Bu sorular en temel olanından ("Hayvanlar, bilimsel araştırmalarda kullanılmalı mıdır?") diğer önemli sorulara (hayvanların bakımı ve korunması hakkındaki sorulara) kadar geniş bir yelpaze çizmektedir. APA'nın Etik Yönetmeliği'ne göre açık şekilde, *etik açıdan araştırmacılar, hayvanların edinilmesi, bakımı, kullanımı ve araştırmanın sonunda federal yasalara ve mahallî yasalara ve düzenlemelere, ve profesyonel standartlara uygun şekilde yok edilmeleri hakkında bilgi edinmek zorundadır.* Kısmen de 1980'lerdeki hayvan hakları savunucularının belirttikleri endişelere cevap olarak, araştırmacıların federal, mahallî ve devlet mevzuatının birçok talebini yerine getirmeleri gerekir; bu taleplere, ABD Tarım Bakanlığı veterinerlerinin, hayvanların kaldıkları yerleri denetlemeleri de dâhildir (National Academy of Science, 2011). Bu düzenlemeler bilim camiası tarafından ve laboratuvar hayvanlarının korunmasını savunan gruplara üye hayvan araştırmacıları tarafından çoğunlukla olumlu karşılanır. APA, psikoloji araştırmalarında hayvanlar kullanılacağı zaman uyulması gereken bir prensipler listesi geliştirdi. Bu prensipler, APA'nın Hayvan Araştırmaları ve Etik Komitesi (CARE) tarafından desteklenen sitede bulunabilir: http://www.apa.org/science/leadership(care/index.aspx.

Hayvanlarla araştırma, temel amacı araştırma hayvanlarının sağlığını korumak olan çok denetimli bir alandır. Sadece araştırma yapma ve özel türlere bakım verme niteliklerine sahip bireylerin hayvanlarla çalışmasına izin verilmelidir. Sadece alternatif işlemler yoksa ve bilimsel, eğitsel veya uygulamadaki amaçlar yapılacak işlemleri haklı çıkarıyorsa, hayvanların acı ve rahatsızlık çekmelerine yol açan işlemler kullanılabilir. Daha önce de belirttiğimiz gibi,

hayvan araştırmalarını denetleme kurulu (animal review board - IACUCs) ABD kamu sağlığı hizmetlerinden parasal destek alan araştırma tesislerini denetlemektedir. Bu komiteler hayvanların çektiği ağrının kontrolü, hayvanın hayatına son verme, barındırma ve personelin eğitimi konularında uygulanan işlemlerin yeterli olup olmadığını belirlemektedir. IACUCs ayrıca, yapılmakta olan deney(ler)in yeni ve önemli bilgiler elde etmek için yeterli olup olmadığını ve bir hayvan modeli kullanmanın mı yoksa başka bir model (örneğin, bir bilgisayar simülasyonu) kullanmanın mı daha uygun olduğunu da inceler ve denetler (Holden, 1987; ayrıca bk. Ulusal Bilimler Akademisi, 2011).

Bununla birlikte, hassas herhangi bir etik meselede olduğu gibi, araştırmalarda hayvan kullanımı konusunda da uzlaşmalar gereklidir. Örneğin, hayvan araştırmalarına alternatifler bulununcaya kadar, insan hastalıklarıyla ve ağrılarıyla savaşmak için araştırmalarda hayvan kullanımı ihtiyacı ile bu amaçla kullanılan hayvanların sağlığını koruma arasında bir denge kurulmalıdır (Goodall, 1987). Ancak, hayvanların kullanımının, araştırma sonuçlarının uygulanması yakın bir gelecekte mümkün olmadığı durumlarda bile sınırlanmaması önemli bir noktadır (Fowler, 1997). Bazılarına göre araştırmalarda hayvan kullanımının bir değeri yoktur, çünkü sonuçların hemen insanlara uygulanabileceği açık değildir. Böyle bir yaklaşım bizi, tüm temel araştırmaların da geçersiz olduğu sonucuna götürür; bu anlayış, hem araştırma hem de uygulama alanında tüm psikolojinin temellerini tehdit eder (Fowler, 1997).

Çoğu bilim insanı hayvanlara gereksiz yere acı çektirilmemesi gerektiği hususunda hemfikir olmakla birlikte, bu insanlar bürokratik işlemlerin ağırlaşmasına ve araştırmayı zorlaştıran maliyetlere de karşıdırlar. Ağır kısıtlamalar ve yüksek maliyetlerle birlikte, hayvan hakları savunucuları arasındaki aşırı grupların araştırmacılara ve araştırma kurumlarına yönelttikleri olumsuz tanıtım (ve duygusal gösteriler), genç bilim insanlarını hayvan araştırmaları alanına girmekten caydırabilir (Feeney, 1987). Eğer böyle olursa hâlen tedavisi olmayan hastalar ve daimî felçliler, araştırmacıların çare bulabileceklerine dair ümitlerini kaybederler. Açık şekilde, hayvan araştırmalarının insanın durumuyla bağlantısına dair birçok tartışmayı çevreleyen meseleler karmaşıktır (bk. Kutu 3.3) ve mantıklı ve dengeli olmayı gerektirmektedir (Ulrich, 1992).

PSİKOLOJİ ARAŞTIRMALARININ RAPOR EDİLMESİ

- Araştırmacılar, yaptıkları araştırmanın sonuçlarını bilimsel dergilerde kendi meslektaşlarına duyurmak isterler ve APA'nın Etik Yönetmeliği bu sürece ilişkin prensipleri belirler.
- Yayında kimlerin yazarlık payesi alacağına dair kararlar, kişilerin yaptıkları bilimsel katkıların önemine dayanır.
- Araştırmanın etik prensiplere uygun şekilde bildirilmesi, aynı alanda başkalarının yaptıkları araştırmalara uygun şekilde referans verilmesini gerektirir; böyle yapılmazsa, aşırmacılık (plagiarism) yapılmış olur.
- Uygun şekilde atıfta bulunma, materyal başka bir kaynaktan aynen alınıyorsa tırnak içinde verilerek ve eğer orijinal bir kaynak kullanılmıyorsa ikinci kaynaklar belirtilerek yapılır.

Tamamlanmış bir araştırma, araştırmacılar bu araştırmanın raporunu psikoloji alanında yayımlanan düzinelerle bilimsel dergiden birisine gönderdikleri zaman, bilimsel literatürün bir parçası olma yolculuğuna başlamış olur (yayım

KUTU 3.3

İNSAN VE HAYVANIN AHLÂKÎ STATÜSÜ

Etik kararlar vermek, çoğunlukla, birbirine karşıt olan felsefî yaklaşımlara neden olur. Bu durum, hayvanların araştırmalarda kullanılması üzerine yapılan tartışmalarda açıkça görülür. Bu tartışmanın odağındaki mesele ise insanın ve hayvanın "ahlâkî statüsü"dür. Filozof Peter Singer'ın da (1990, s. 9) işaret ettiği gibi, genel olarak kabul edilmiş olan iki ahlâk prensibi var:

1 Tüm insanlar ahlâken eşittir.
2 Tüm insanlar ahlâken diğer hayvanlardan üstündür.

Ve Singer devam ediyor: "Bu prensiplere dayanarak insanın sağlığını, hayvanların acı çekmelerinden üstün tutmamız gerektiği, yaygın şekilde savunulmaktadır. Bu varsayımı hayatımızın birçok alanında görebiliriz: Çiftçilikte, avcılıkta, deneylerde ve eğlencede" (s. 9).

Ama Singer'ın kendisi bu yaygın görüşe katılmıyor. Ona göre "İnsanın acısını hayvanın acısından önemli görmemizin akla dayanan etik bir gerekçesi yoktur." (s. 9). Singer'a göre dinî görüşe başvurmadığımız sürece (ki Singer çoğulcu bir toplumda karar vermek için dinin esas alınmasını reddeder), "insan olmanın" özel bir ahlâkî statüsü olamaz. Bu görüşün kökleri David Hume (1711-1776), Jeremy Bentham (1748-1832) ve John Stuart Mill'in (1806-1832) savundukları ve faydacılık (utilitiarianism) olarak bilinen görüştür. Buna göre, her ne zaman alternatif eylemler arasında bir seçme şansımız varsa, işin içinde olan herkes için en iyi sonuçları (en çok mutluluk) vereni seçmeliyiz. Mesele, söz konusu bireyin mutluluğu ve mutsuzluğu, zevki ve acıyı deneyimleme kapasitesine sahip olup olmadığıdır. "Ağrı ve acı kötüdür ve ırk, cinsiyet veya tür farkı gözetmeksizin engellenmelidir." (Singer, 2011, s. 53).

İnsanın ve hayvanın ahlâkî statüleri ve bunun psikoloji araştırmalarıyla ilişkisi hakkında siz neler düşünüyorsunuz?

süreci hakkında detaylı bilgi için, bk. Bölüm, 13). Bir araştırmayı psikoloji dergilerinde yayımlamanın temel amacı, araştırmanın bulgularını bilim camiasına ve genel olarak topluma duyurmaktır. Bu yayın aynı zamanda, araştırmacının ve araştırmayı destekleyen kurumun adını duyurmanın da bir yoludur. Ama bilimsel bir araştırmayı yayımlatmak her zaman kolay bir süreç değildir; özellikle de araştırmacı araştırmasını prestijli bir bilimsel dergide yayımlatmak istediği zaman durum daha zordur. Psikoloji biliminde araştırma yayımlamanın önemi yüzünden, APA'nın Etik Yönetmeliği'nde bu süreçle ilgili prensipler yer almaktadır.

Bir bilimsel araştırmanın sonuçlarını yayımlamanın etik standartları, Etik Yönetmeliğin tartışmış olduğumuz diğer alanlarına kıyasla daha açıktır. Ama bu alanda bile yayınların yazarlarını ve sıralarını belirleme ve aşırmacılık gibi konularda karar vermek her zaman kolay değildir. Bir araştırmaya çoğunlukla birden fazla insan katılır. Meslektaşlar araştırmanın deseni için önerilerde bulunur; öğrenciler veya mezunlar katılımcıların test edilmesi ve verilerin organize edilmesi için yardımcı olur; teknisyenler araç-gerecin kurulmasını ve işlemesini sağlar, ve uzman danışmanlar istatistiksel analizler için danışmanlık yapar. Yayım için sunulacak metin hazırlanırken bu insanların hepsi araştırmanın "yazarları" olarak mı kabul edilecek? *Yazarların belirlenmesi* (publication credit), araştırma projesine önemli katkısı olan kişileri yazar olarak belirleme sürecine denir. Yayımlanmış bir bilimsel araştırmanın yazarlığı, çoğunlukla bir kimsenin

bilim alanındaki yeteneğinin ve motivasyonunun ölçütü olarak kullanıldığı için, *bir projeye katkıda bulunmuş olanların dürüstçe belirtilmesi önemli bir husustur.*

Birisinin bir araştırma projesine katkısının, o araştırmaya ait makalede "yazar" olarak mı yoksa daha önemsiz bir şekilde mi (örneğin, dip not şeklinde) gösterilmesi gerektiğine karar vermek her zaman kolay değildir. Yazarlık verilecekse, bu defa kaçıncı yazar olarak belirtileceğine de karar verilmesi gerekir. Çok yazarlı bir makalenin "ilk yazar"ı olmak, genellikle, ikinci yazardan daha fazla katkısı olduğu (ikincinin de üçüncüden. ...) anlamına gelir. Kaçıncı yazar olunacağına karar verilirken, katkının bilimsel önemi (örneğin, kavramsal yönden araştırmaya yapılan katkılar) esas alınmalıdır; araştırmanın yapılması için harcanan zaman ve enerji değil (bk. Fine ve Kurdek, 1993).

Yazarlık sırasını belirlemeye ilişkin etik endişeler, birçok şekilde ortaya çıkabilir. Örneğin, bir öğretim üyesinin bir öğrencinin yaptığı çalışmada kendisini yazar olarak göstermesi etik dışı olduğu gibi, öğrencilere hak etmedikleri bir yazarlık unvanı verilmesi de etik değildir. Bu sonraki durum örneğin, elemenin yüksek olduğu bir lisansüstü programına giriş sınavında öğrenciye üstünlük sağlamak istenildiği bir durumda ortaya çıkabilir. Fine ve Kurdek'e (1993) göre, öğrencilere hak etmedikleri bir yazarlık payesi vermek, yanlış yere onun uzman olarak görülmesine, emsalleri üzerinde haksız bir avantaj sağlamasına ve belki de bu öğrenciden gerçekçi olmayan bir beklenti oluşmasına yol açabilir. Yazarlık unvanının kime verilmesi gerektiği konusunda öğretim üyeleri ve öğrenciler işbirliği yapmalı ve araştırmanın başlarında, hangi düzeyde katkının yazarlık hakkı vereceği konuşulmalıdır. Öğretim üyeleri ile öğrenciler arasındaki güç ve mevki farkı nedeniyle, bu konudaki tartışmayı öğretim üyelerinin başlatmaları gerekir (bakınız, Behnke, 2003).

Temel Kavram

Sadece profesyoneller için değil aynı zamanda öğrenciler için de araştırmanın rapor edilmesinde çok sıkıntılı bir konu, **aşırmacılık**tır (plagiarism). Bir defa daha, etik standardın açık olduğunu belirtelim: Bir başkasına ait bir çalışmanın önemli bir kısmını veya unsurlarını, kendinize aitmiş gibi yayımlamayın. Ama "önemli bir kısmını veya unsurlarını" teşkil eden şey nedir ve bir araştırmacı, bir başkasına ait bir çalışmayı kendisine aitmiş izlenimi vermekten nasıl kaçınabilir? Bu konuda karar vermek, ip üstünde cambazlık yapmaya benzer. Bir yanda çalıştığı alana anlamlı bir katkıda bulunarak tanınma amacı var; öte yanda ise daha önceden bu alana yapılmış olan katkıların kabul edilmesi var. Hem profesyonellerin hem de öğrencilerin aşırmacılık yapmaları olgusu şunu düşündürmektedir: Birçok insan çoğunlukla, başkalarının çalışmalarını övmektense, kendilerinin tanınmasını tercih ederek ipin üzerindeki hareketlerini buna göre belirlemektedir.

Bazen aşırmacılık, özensizlikten kaynaklanabilir (örneğin, bir düşüncenin bir başkasına ait olmadığından emin olmak için bir kaynağı ikinci defa kontrol etmeyi ihmal etmek gibi). Bu tip hatalar da yine aşırmacılıktır; *ihmal, haklı bir gerekçe değildir.* Çok kolay hata yapılabilir. Örneğin, araştırmacılar (ve öğrenciler) bazen bir paragrafın "ne kadarının" tırnak işareti kullanılmadan ya da kaynağı bir başka şekilde belirtilmeden kullanılabileceğini sorarlar. Eğer bir başkasına ait bir düşüncenin temel bir kavramını temsil ediyorsa, bir tek kelime veya kısa bir cümlecik bile çok önemli hâle gelir. Ne kadar materyalin bir çalışmanın temel bir ögesi olduğunu belirten açık bir kural olmadığı için, öğrenciler özellikle başkalarının çalışmalarına atıfta bulunurken çok dikkatli olmalıdır. Bazen, özellikle öğrenciler, bir kaynaktan doğrudan yaptıkları bir alıntıyı tırnak içine almadıkları için aşırmacılık yapmış olurlar. *Her ne zaman bir materyal orijinal kaynağından doğrudan alınırsa tırnak içine*

TABLO 3.1 DOĞRU ALINTILAMA VE AŞIRMACILIK ÖRNEKLERİ

Asıl Metin (Doğru Yapılmış Doğrudan Alıntılama Örneği)
"İçtihat hukukundaki gelişmeler hakkında bilgi alan polis, çeşitli sorgulama yöntemleri kullanır – bunlar arasında, zanlıya sahte kanıtların sunulması (örneğin, sahte yalan makinesi, sahte parmak izi, veya sahte tanık teşhisleri gibi başka sahte adlî test sonuçları), Tanrıdan ve dinden yardım istemek, dostmuş gibi görünmek ve cezaevi muhbiri kullanmak gibi yöntemler vardır" (Kassin ve Kiechel, 1996, s. 125).

Aşırmacılık Örneği (Yeniden İfade Edilmiş Materyal için Kaynak Belirtilmemiş)
Suç itiraf ettirmek için kullanılan aldatıcı sorgulama yöntemleri üzerine yapılan araştırmalar önemlidir, çünkü polis zanlıları sorgularken sahte kanıtlar (örneğin, sahte test sonuçları gibi) ve sahte tanıklar kullanır. Sorgulamayı yapanlar aynı zamanda, zanlılara dostmuş gibi yaklaşarak da baskı yaparlar.

Yeniden İfade Edilmiş ve Doğru Şekilde Kaynak Gösterilmiş Materyal Örneği
Zanlıların suçlarını itiraf etmelerini sağlamak için aldatıcı sorgulama yöntemleri kullanılması üzerine yapılan araştırmalar önemlidir, çünkü polis zanlıları sorgularken sahte kanıtlar (örneğin, sahte test sonuçları) ve sahte tanıklar kullanmaktadır (Kassin ve Kiechel, 1996). Ayrıca, Kassin ve Kiechel'e göre sorgulamayı yapanlar, zanlının dostuymuş gibi davranarak, itirafta bulunması için baskı yaparlar.

Zechmeister, Zechmeister ve Shaughnessy'den (2001, s. 71) alınmıştır.

alınmalı ve kaynak uygun şekilde belirtilmelidir. Bir materyali aynen değil de kendi cümlelerinizle yeniden yazdığınız (paraphrase) zaman da o materyalin kaynağını belirtmeniz gerekir. Etik prensip şöyledir: *İfadeleri aynen kullandığınız zaman ve kendi cümlelerinizle yeniden ifade ettiğiniz zaman düşüncelerinizin kaynağını belirtmeniz gerekir*. Doğru ve yanlış alıntılara birkaç örnek, Tablo 3.1'dedir.

Aşırmacılık, araştırmacılar ikinci kaynakları belirtmedikleri zaman da meydana gelir. *İkinci kaynak* demek, başka (original) kaynağı tartışan kaynak demektir. İkinci kaynaklar, kitaplar ve *Psychological Bulletin* gibi bilimsel dergilerde yayımlanmış inceleme yazıları olabilir. Elinizdeki bir düşüncenin veya bulguların tek kaynağı bir ikinci kaynak ise, *o düşünceyi veya bulguları sanki orijinal kaynağı okumuşsunuz gibi bildirmeniz* etik dışıdır. İkinci kaynağa atıfta bulunmaktansa, orijinal kaynağı bulmaya ve okumaya çalışmak çok daha iyidir. Eğer bu mümkün değilse, orijinal çalışmaya atıfta bulunacağınız zaman" ...in naklettiği gibi ..." benzeri bir ifade kullanarak okuyucuya bilgi vermeniz gerekir. İkinci kaynağa atıfta bulunarak okuyucuya diyorsunuz ki "Benim orijinal kaynak hakkında sunduğum bu yorum, bir başkasına aittir." Bir defa daha ortaya çıkıyor ki kaynakçayı doğru şekilde göstermeyi ihmal etmek kabul edilebilir bir hata değildir ve bazı şanssız durumlarda araştırmacılar –öğretim üyesi veya öğrenci– kariyerlerinin aşırmacılık suçlamalarıyla mahvolduğuna tanık oldular.

ETİK AÇIDAN UYGUN BİR ARAŞTIRMA YAPMANIN AŞAMALARI

- Etik karar verme, yapılmak istenilen araştırma ortamına ait olguların incelenmesini, konuyla bağlantılı etik meselelerin ve prensiplerin belirlenmesini ve çeşitli bakış açılarının ve alternatif yöntem ve işlemlerin ele alınmasını kapsar.
- APA'nın bir dergisine bir araştırma sunan yazarların, yaptıkları araştırmanın etik standartlara uygunluğunu gösteren belgeleri de sunmaları gerekir.

KUTU 3.4

ETİK AÇIDAN TARTIŞMALI BİR ARAŞTIRMAYI ETİK ŞEKİLDE YAPMAK

Stanley Milgram'ın 1960'lardaki ünlü deneyleri (Kutu 3.2'de betimlendi), açık bir soru sordu (Blass, 2009; Milgram, 1963, 1965, 1974): İnsanlar hangi koşullar altında, kendilerinden başka insanlara acı çektirmelerini isteyen bir yetkiliye itaat ederler?

Aldatma ve risk konusundaki endişeler nedeniyle, on yıllardan beri Milgram'ın işlemlerini kullanan araştırmalar yapılmadı ta ki Santa Clara Üniversitesinden Jerry M. Burger Milgram'ın deneylerini tekrarlamaya çalışana kadar (Burger, 2009). Onun sorusu şuydu: İnsanlar bugün hâlâ aynı şekilde itaat ederler mi?

O hâlde Burger'ın sorduğu soru şudur: Milgram'ın deneylerini, etik açıdan kabul edilebilir bir şekilde nasıl yapabilirsiniz? Bu sorunun cevabı, Burger'ın yaptığı replikasyonda kullandığı işlemdedir. Kendi kampüsündeki IRB'nin endişelerini gidermek ve katılımcıların sağlığını güvence altına almak için Burger, Milgram'ın orijinal deneyinde aşağıdaki değişiklikleri yaptı:

- Araştırmadan önce, katılımcıları dikkatle eleme.
- Katılımcılara, istedikleri zaman araştırmadan çekilebileceklerini ve yine de kendilerine ödeme yapılacağını açık şekilde söyleme (gerçekte, bütün katılımcılara araştırma başlamadan önce ödeme yapıldı).
- Çok daha yumuşak bir "örnek"şok kullanma.
- Katılımdan hemen sonra katılımcılara, araştırmada aldatma yapıldığını bildirme, araştırmanın amacını açıklama, ve "kurban"a zarar verilmediğini gösterme.
- Deneyi bir klinik psikologa yaptırmak; bu psikolğun görevi, katılımcının aşırı bir stres yaşadığını düşündüğü zaman uygulamaya son vermekti.

Bu replikasyondaki işlemin Milgram'ın uyguladığı işlemden en önemli farkı şuydu: Katılımcılar şok jeneratöründe (sahte) 150 volta çıktıkları zaman (Milgram'da bu 450 volt idi), işlem sona erdirildi. Burger bu yaratıcı değişikliği, orijinal deneylerdeki her beş katılımcıdan dördünün, bir defa 150 voltu uyguladıktan sonra 450 volta kadar çıktıklarını keşfettiği zaman yaptı.

Burger'ın bulguları hâlâ huzursuz etmeye devam ediyor: "İtaat oranları, Milgram'ın 45 yıl önceki bulgularından daha düşüktü ama bu fark çok azdı" (Burger, 2009, Abstract, s. 1).

Burger'ın (2007) bildirdiği önemli bir husus da şudur: "Verileri toplamanın üzerinden bir yıldan fazla bir zaman geçtikten sonra herhangi bir katılımcının, araştırmaya katıldığı için zarar gördüğüne dair bir işaret görmedim. Tersine, katılımcıların hem araştırma sonrası bilgilendirme esnasında hem de sonraki iletişimlerimiz esnasında gösterdikleri ilgiye hep şaşırdım." (s. 17).

Burger'ın araştırması gösteriyor ki insan katılımcıların sağlığı ile önemli psikolojik meselelerin araştırılması arasındaki doğru dengeyi bulmak için bazen karmaşık işlemler gerekiyor (bk. Behnke, 2009).

İnsan davranışı hakkında bilgi edinmek için, katılımcıların ciddi zarar görebilecekleri risklere atılmaları gerekir mi? Psikologların katılımcıları araştırma amacıyla aldatmaları gerekir mi? Araştırma amacıyla hayvanların acı çekmelerine izin verilmesi kabul edilebilir bir tutum mudur? Etik karar vermenin bir parçası olan bu soruların cevaplanması zordur ve incelikli karar verme süreçleri sonunda verilecek cevaplar, herkesi "mutlu" etmeyecektir. Etik kurallara uygun bir bilgilendirilmiş onam sürecinde, şu aşamalar izlenmelidir:

1. Önerilen araştırma koşullarını (örneğin, katılımcılar, işlem vb.) inceleyin.
2. Araştırmayla bağlantılı etik meseleleri, yönetmelikleri ve yasaları belirleyin.
3. Farklı görüş açıları üzerinde düşünün (örneğin; katılımcılar, araştırmacılar, kurumlar, toplum, ahlâkî değerler gibi).
4. Önerilen araştırmayı yapmama seçeneği dâhil, alternatif yöntemler veya işlemler ve bunların sonuçları üzerinde düşünün.

Bu faktörleri dikkatle hesaba katarak önerilen araştırma hakkında "doğru" bir karar vermek, araştırmanın ve etik meselelerin dikkatli bir incelemesiyle mümkündür; sadece araştırmacıları veya araştırmanın içindeki başka kimseleri neyin "mutlu" edeceğine bağlı değildir. Kutu 3.4, Milgram'ın itaat işlemlerini kullanan yeni araştırmayı betimlemektedir.

APA dergilerinden birisine sunulmuş olan bir araştırmanın yazarları, araştırmanın etik standartlara uygun olduğunu gösteren belgeleri de sunmalıdır (bk. *Publication Manuel of the American Psychological Association*, APA, 2010b). Bu formlar *Publication Manuel*'de (s. 233-235) ve APA journal'ın web sitesinde (http://www.apa.org/pubs/journals) bulunabilir. Elbette etik meseleler, araştırmaya başlamadan önce, araştırma sürecinin kendisi devam ederken, problemler (örneğin, katılımcıların beklenmedik tepkileri) ortaya çıktıkça, ve taslak bir dergiye sunulduğu zaman editörler ve hakemlerle tartışmaya hazırlanırken göz önüne alınmalıdır. Araştırma boyunca etik uygunluğun güvence altına alınmasına yardım etmek için APA, bir Etik Uygunluk Kontrol Listesi yayımladı (bk. *Publication Manual*, s. 20). Kontrol listesinde, bu bölümde tartışılmış olan birçok mesele yer almaktadır; bunlara kurumsal inceleme, bilgilendirilmiş onam, hayvan katılımcılara yapılacak muameler (eğer kullanılıyorsa), yayımlanmış başka araştırmalara uygun şekilde atıfta bulunulması ve yazarların isimlerinin sıralanması dâhildir. Unutmayın: APA'nın uygunluk formlarında açıklanmış olan bu ve başka meseleleri, araştırmanıza başlamadan *önce* dikkatle incelemeniz gerekir.

ÖZET

Psikoloji araştırmalarında birçok etik sorun ortaya çıkar. Bu nedenle, bir araştırmaya başlamadan önce hem APA'nın Etik Yönetmeliği'ndeki ve yasalardaki özel etik meseleleri hem de araştırmanızla bağlantılı düzenlemeleri incelemelisiniz. Çoğu durumda, araştırmaya başlamadan önce resmî kurumsal onay –örneğin, bir IRB'den veya IACUC'den– alınması gerekir. IRB'nin bir işlevi, önerilen araştırmanın risk/fayda oranı hakkında bir uzlaşmaya varmaktır. Risk fiziksel, psikolojik veya sosyal zarar şeklinde olabilir. İnsan katılımcılarla yapılan araştırmaların çoğunda, bilgilendirilmiş onam alınmalıdır. En az düzeyden daha fazla risk olduğu durumlarda araştırmacılar, katılımcıları korumak için ek önlemler almalıdır ve araştırma sonrası bilgilendirme yapmalıdır. Araştırmacılar katılımcılardan bilgi sakladıkları zaman veya araştırmanın mahiyeti hakkında yanlış bilgi verdikleri zaman ciddi etik sorunlar ortaya çıkar. Araştırmada aldatma yapıldığı zaman, araştırma sonrası bilgilendirme sırasında, neden aldatma yapıldığı açıklanmalıdır. Araştırma sonrası bilgilendirme aynı zamanda, katılımcıların kendilerini tam olarak araştırmanın içinde hissetmelerine ve onların deneydeki işlemleri ve görevlerini nasıl algıladıkları hakkında deneyciye bilgi sağlar. İnternette araştırma yapmak, bir araştırmacının karşısına yeni etik sorunlar çıkarır; araştırmacıyı, böyle bir araştırma yapmadan önce IRB üyelerine ve internette veri toplama konusunda deneyimli başka araştırmacılara danışmaya zorlar.

Hayvanları test eden psikologların, federal ve devlet yönetmeliklerine uymaları ve genel olarak, hayvanların sağlığını korumaları gerekir. Ancak bazı koşullar altında hayvanların acı çekmelerine izin verilebilir: Örneğin, başka bir yöntem olmadığı zaman, araştırma bulgularının bilimsel, eğitsel veya somut bazı sorunların çözümüne katkılarının çok değerli olduğu kabul edildiği zaman

olduğu gibi. Birçok insan hastalıklarla savaşmak için, hayvan araştırmalarında kullanılabilecek alternatif yöntemler bulununcaya kadar, onların sağlığını koruyarak laboratuvarlarda kullanılmaları gerektiğini düşünmektedir.

Psikoloji alanındaki araştırma bulguları rapor edilirken, araştırmaya katılanların makaledeki sıralamaları, onların araştırmaya yaptıkları katkı ile uyumlu olmalıdır. Eğer bir araştırmacı düşüncelerini geliştirirken daha önceki bir araştırmadan faydalanmışsa, önceki araştırmanın sahiplerini belirtmelidir. Bunun yapılmaması, ciddi bir etik ihlaldir: Aşırmacılık (plagiarism). Etik karar vermek için önerilen araştırmanın koşullarını incelemek, o araştırmayla ilişkili etik meseleleri ve yönetmelikleri belirlemek ve farklı bakış açılarını ve alternatif yöntemleri ve işlemleri göz önüne alarak üzerinde düşünmek gerekir. Yayınlanması için bir araştırmayı APA'nın bir dergisine gönderen araştırmacılar, aynı zamanda, yaptıkları araştırmanın etik kurallara uygunluğunu gösteren belgeleri de araştırmalarına eklemeleri gerekir.

TEMEL KAVRAMLAR

risk/fayda oranı 62
minimal risk 65
bilgilendirilmiş onam 67
mahremiyet 71
aldatma 73
araştırma sonrası bilgilendirme 77
aşırmacılık 82

GÖZDEN GEÇİRME SORULARI

1 Araştırmacılar neden araştırmalarına başlamadan önce, araştırma önerilerini Kurumsal İnceleme Kurulları'na (Institutional Review Boards - IRBs) veya Kurumsal Hayvan Bakım ve Kullanım Kurulları'na (Institutional Animal Care and Use Committees - IACUCs) gönderiyorlar ve bu kurulların araştırma sürecindeki işlevleri nedir? Açıklayın.
2 Etik kararlar verilirken, risk/fayda oranı nasıl kullanılır? Bir araştırmanın potansiyel faydaları belirlenirken hangi faktörler etkili olur?
3 Riski olmayan bir araştırma neden mümkün değildir? Araştırmacı, katılımcıların "risk altında" olup olmadıklarını belirlemek için hangi standartları kullanır? Katılımcıların özellikleri, risk değerlendirmesini nasıl etkiler?
4 Psikoloji araştırmalarında ortaya çıkabilecek üç risk tipini belirtin: Fiziksel risk, psikolojik risk ve sosyal risk. Araştırmacılar, sosyal riskin ortaya çıkmaması için hangi önlemleri alırlar?
5 İnternette araştırma yapmanın en önemli üç etik riski hangileridir?
6 Araştırmacı, katılımcıdan bilgilendirilmiş onam almak için, etik açıdan hangi bilgileri katılımcıya vermek zorundadır?
7 Bir bilginin özel mi yoksa umumî mi olduğuna karar verirken, araştırmacının göz önüne alması gereken üç husus hangileridir?
8 Psikoloji araştırmalarında neden bazen aldatma yapmak gerekir? Araştırmacı, aldatmaya başvurmadan önce hangi soruları cevaplamalıdır? Katılımcıları aldatmak, hangi koşullar altında daima etik dışı bir yöntemdir?
9 Araştırma sonrası bilgilendirme, ne şekilde yapılırsa katılımcılar için, ne şekilde yapılırsa araştırmacı için faydalı olur?
10 APA'nın Etik Yönetmeliği'nde, araştırmalarında hayvan kullanan araştırmacılar için hangi etik zorunluluklar getirilmiştir?

11 APA Etik Yönetmeliği2ne göre, araştırmalarda hayvanlara acı veya stres yaşatmak için hangi koşulların yerine getirilmesi gerekir ?
12 Araştırmacılar, birisinin bir araştırmada yazar olup olmayacağına nasıl karar veriyorlar?
13 Bir yazarın, orijinal bir kaynaktan veya ikinci bir kaynaktan alıntı yaptığı zaman aşırmacı durumuna düşmemesi için uyması gereken işlemler hangileridir?
14 Önerilen bir araştırmanın yapılıp yapılmamasına karar vermek için uyulması gereken işlemler, etik açıdan hangi aşamalardan geçer?
15 Araştırmacılar, araştırmalarını APA'nın bir dergisine gönderdikleri zaman hangi belgeleri eklemeleri gerekir?

DÜŞÜNDÜRÜCÜ SORULAR

Not: Diğer bölümlerden farklı olarak, bu bölümdeki Düşündürücü sorular'a veya İleri Alıştırmalar'a cevap vermeniz gerekmiyor. Etik ikilemleri çözmeniz için uygun etik standartları uygulayabilmeniz ve deneyimleri ve bilgileri sizinkinden farklı olan başka insanlarla tartıştıktan sonra, önerilen araştırma hakkında bu insanlarla uzlaşmanız gerekiyor. Bu nedenle, sizinkinden farklı görüş açılarını hesaba katmanız gerekiyor. Bu önemli meselelere, bir grup tartışmasının önemli bir parçası olarak yaklaşmanızı öneriyoruz.
Bu bölümdeki ilk iki zorlu soru, gerekçesi (rationale) ve yöntemi gerçek bir araştırmanınkine benzeyen hipotetik bir araştırmaya aittir. Bu soruları cevaplamak için, APA'nın Etik Yönetmeliği'ni ve bu bölümde bahsettiğimiz etik kararlara varmanız için gerekli diğer materyali bilmeniz gerekir. Ayrıca, bu bölümün sonunda özetlediğimiz etik karar verme aşamalarını da bilmeniz gerekir. Göreceğiniz gibi sizin göreviniz, belirli etik prensiplerin ihlal edilip edilmediğine karar vermek ve yapılmak istenilen araştırma hakkında önerilerde bulunmaktır; tabi en önemlisi de araştırmacının bu araştırmaya devam etmesine izin verilip verilmemesi gerektiğine karar vermektir.

1 Kurumsal İnceleme Kuruluna (IRB) Gelen Araştırma Önerisi

İstenilenler Bir kurumsal inceleme kurulunda üye olduğunuzu varsayın. Kurulda sizden başka bir klinik psikolog, bir sosyal psikolog, bir sosyal hizmet uzmanı, bir felsefeci, bir din adamı, bir tarih profesörü ve bir iş yeri yöneticisi var. Kurulun incelemesi için sunulmuş olan araştırma önerisi aşağıdadır. Araştırmacıya hangi soruların sorulmasını istersiniz? Sunulduğu şekliyle bu araştırmanızın kendi kurumunuzda yapılmasına izin verip vermeyeceğinize ve onay vermeden önce araştırmada değişiklik yapılması gerekip gerekmediğine veya önerinin onaylanıp onaylanmaması gerektiğine karar verin (Kurula onay için sunulmuş olan gerçek araştırma önerisinde, aşağıda belirttiklerimizden daha ayrıntılı bilgiler sunulmuştur.).

Araştırmanın Gerekçesi Kadın ve erkek (daha çok kadın) üniversite öğrencileri, her hafta saatlerce internet (özellikle de Facebook) sitelerinde zaman geçirmektedirler. Önerdiğimiz araştırma, beş kişilik özelliğini belirleyen Büyük Beş Kişilik Testi'ni kullanarak, bilgisayardaki davranışın yordayıcılarını (predictor) araştıracaktır. Bu özellikler şunlardır: Deneyime açıklık, dürüstlük, dışa dönüklük, kabul edilirlik ve nevrotiklik (duygusal istikrar). Bu iyi bilinen Beş Faktör Kişilik Modeli iş performansı, psikolojik mutluluk ve akademik başarı gibi çok sayıda sonucu yordamada kullanıldı. Yakın zamanlarda bu beş kişilik özelliği, Facebook'a giriş sayısı, Twitter kullanıcılarının tipi gibi hususlarda sosyal medya kullanımını yordamak için kullanıldı.

Büyük Beş Kişilik Testi'nin bir şekli, bir beşli ölçek kullandı; bu ölçek üzerinde katılımcıların, kendileri hakkında belirtilen ifadelere ne ölçüde katıldıklarını belirtmeleri istendi. Beş boyutla ilişkili örnek maddeler şunlardır:

- Keskin bir hayal gücüm vardır (Açıklık)
- Programlarıma uyarım (Dürüstlük)
- Hayat insanıyım (Dışa dönüklülük)
- İnsanlarla ilgileniyorum (Kabul edilirlik)
- Kolayca sinirlenirim (Nevrotiklik)

Önerilen araştırma Büyük Beş kişilik özelliğinin, üniversite öğrencilerinin Facebook sayfaları kullanılarak değerlendirilip değerlendirilemeyeceğini araştıracaktır. Özel

(devam ediyor)

olarak, üniversite mezunu olan iki hakem, her bir kullanıcının kişiliğini sınıflamak için bir kimsenin Facebook sayfasından alınan bilgilerini (örneğin; profili, geçmişi, sevdikleri gibi) kullanacaktır.

Yöntem Katılımcılar, büyük bir psikolojiye giriş sınıfından alınacaktır ve asıl araştırmacı, bu sınıfın öğretmenidir. Sınıfın Facebook kullanan öğrencilerinden, dersin öğretmenini Facebook'ta arkadaş olarak seçmeleri istenecek; böylece öğretmenin asistanları bu öğrencilerle iletişim kurabilecektir (örneğin; soruları cevaplama, dersle ilgili mesajları verme, geri bildirim verme, derse ilişkin duyuruları yapma gibi). İki mezun asistan, öğretmenin açmış olduğu bu Facebook hesabını kullanarak, öğrencilerin Facebook sayfalarını sınıfta izleyecektir. Bu izleme, 8 hafta boyunca haftada iki defa öğrencilerin sayfaları "ziyaret" edilerek yapılacaktır. Her ziyarette iki mezun asistan, sayfada bulunan bilgileri öğrencileri Beş Büyük boyut üzerinde değerlendirmek için kullanacaktır. Bu dereceleme 8 hafta boyunca, yeni bilgiler geldikçe, yenilecektir.

Yarı yılın sonunda bu sınıftaki öğrencilere, Büyük Beş Kişilik Testi'nin kâğıt-kalem versiyonu uygulanacaktır; bu uygulama kişilik psikolojisi dersinin bir parçası olarak yapılacaktır. Öğrenciler bu testin üzerine isimlerini yazacaklar; böylece öğretmen, testin sonuçlarını kendilerine bildirecektir; öğrenciler de Beş Büyük Kişilik Testi üzerine sınıfta yapılan tartışmaya dayanarak, bu sonuçları yorumlayacaklar.

Sınıf içi kişilik testi sonuçları, mezun asistanların Facebook sayfasındaki bilgilerden elde ettikleri sonuçlarla kıyaslanacaktır. Veri analizi, Facebook'tan elde edilen bilgiler kullanılarak, varsa, hangi özelliklerin geçerli şekilde değerlendirilebileceğini belirleyecektir. Bu analiz ayrıca, her iki kişilik derecelemesinin, öğrencilerin sınıftaki derecelerini yordayıp yordamadığını da gösterecektir. İki mezun asistanın Facebook sayfalarını kullanarak yaptıkları değerlendirme ile öğrencilerin sınıftaki test sonuçları arasında pozitif bir korelasyon çıkarsa, Facebook sayfalarından elde edilen bilginin, kişiliği değerlendirmek için kullanılabileceği sonucuna varılacaktır. Ve ayrıca, eğer bu korelasyon bulunursa, kişilik özellikleri ile sosyal medyanın kullanılması arasındaki ilişkiyi incelemek için ek analizler yapılacaktır.

2 Kurumsal Hayvan Bakım ve Kullanım Kuruluna (IACUC) Gelen Araştırma Önerisi

İstenilenler Varsayın ki siz bir Kurumsal Hayvan Bakım ve Kullanım Kurulunda üyesiniz. Sizden başka bu kurulda bir veteriner, bir biyolog, bir felsefeci ve bir iş yeri yöneticisi var. Kurulunuzun incelemesi için gelen bir araştırma önerisinin özeti aşağıdadır. Araştırmacılara hangi soruları sormak istersiniz? Bu şekliyle bu araştırmanın, yöneticisi olduğunuz kurumda yapılmasını ister miydiniz? Onaylanmadan önce değişiklikler yapılması gerekiyor mu? Veya, bu önerinin onaylanmaması mı gerekiyor? (IACUC'a sunulmuş olan gerçek araştırma önerisi, aşağıdakinden daha ayrıntılıdır).

Araştırmanın Gerekçesi Araştırmacıların amacı korteks altı yapıların, limbik sistemde duygu ve saldırganlığı etkileyip etkilemediğini anlamaktır. Bu öneri, aynı laboratuvarda daha önce yapılmış olan bir araştırmaya dayanmaktadır; maymunlar üzerinde yapılan söz konusu araştırmada, beynin çeşitli alt kortikal alanlarında yaratılan hasarlar ile yeme, saldırma ve kur yapma gibi diğer sosyal davranışlardaki değişmeler arasında anlamlı bir ilişki olduğu bulunmuştur. Araştırılacak olan beyin alanları, bazen insanlarda aşırı saldırgan davranışları kontrol altına almak için cerrahî yollarla kesilen alanlardır. Üstelik, önerilen araştırmanın odaklandığı özel korteks altı alan ile aşırı cinsel arzular arasında bir ilişki olduğu görüşü öne sürülmüştür; bu aşırı cinsel arzuları kontrol etmek için insanlara bazen psikolojik tedavi uygulanmaktadır. Önceki araştırmalar, bu davranışları kontrol etmede işe karıştığı düşünülen alanları kesin olarak belirlemede başarısız oldu; önerilen araştırmanın amacı, bu konuda ilerleme sağlamaktır.

Yöntem Araştırmada iki grup maymun kullanılacaktır. Gruplardan birisi (N = 4) kontrol grubu olacak ve sadece görünüşte bir operasyon yapılacaktır: Anestezi verilecek ve kafataslarında bir delik açılacaktır. Daha sonra, deney grubuyla aynı teste tâbi tutulacak ve değerlendirilecektir. Deney grubundaki 4 maymunda ise bir korteks altı yapı olan amigdala bölgesinde küçük bir hasar meydana getirilecektir. Meydana getirilen bu hasarlar, iki hayvanda amigdalanın bir tarafında, diğer iki hayvanda ise amigdalanın diğer tarafında olacaktır. Hayvanlar iyileştikten sonra hepsine çeşitli testler uygulanacaktır. Bu testlerde hayvanların yemek tercihleri, aynı cinsten ve karşıt cinsten maymunlarla olan sosyal ilişkileri ve duygusal tepki verme (örneğin, palyaço maskesi takmış bir deneyci gibi yeni bir korku uyarımına tepki verme) yetenekleri ölçülecektir. Hayvanlar modern bir hayvan laboratuvarında tutulacaktır; ameliyatlar ve sonrasındaki

iyileşme aşaması, uzman bir veteriner tarafından yapılacak ve gözlenecektir. Testlerden sonra hayvanların hayatına son verilecek ve beyinleri, doku incelemesi için hazırlanacaktır (Doku incelemesi, hasarın yerinin ve genişliğinin doğrulanması için gereklidir). Kontrol hayvanlarının hayatlarına son verilmeyecektir; başka deneylerde kullanılmak üzere gruba geri döneceklerdir.

3 Bu bölümde, Milgram'ın ünlü deneyinin (bk. Kutu 3.2) yakınlarda tekrar yapıldığını (replication), ama bu defa katılımcıların korunması için özel önlemler alındığını (Burger, 2009) öğrendik. Aşağıdaki soruları cevaplamadan önce, Milgram'ın orijinal işlemine (bk. Kutu 3.4) konulan sınırlamaları inceleyin.
   A Milgram'ın orijinal deneyinde yapılan aldatmanın, etik açıdan haklı olduğuna inanıyormusunuz? Burger'in tekrarladığı deney için ne düşünüyorsunuz?
   B Psikoloji araştırmalarında aldatmaya izin verilmesi gerektiğine inanıyor musunuz? Neden?

4 Bir yayının yazarlarını belirleme konusunda yaptıkları bir tartışmanın bir parçası olarak, Fine ve Kurdek (1993) bir senaryo hazırlamışlardır. Bu senaryonun aşağıda gördüğünüz bir versiyonu üzerinde düşünün.

Psikoloji bölümünden bir öğrenci, bir öğretim üyesinden, tezinin danışmanlığını üstlenmesini istedi. Öğrenci araştırma konusunu önerdi, ikisi birlikte araştırma yöntemini belirlediler, öğrenci verileri topladı ve sisteme girdi ve öğretim üyesinin yardımıyla verilerin istatistiksel analizlerini yaptı. Öğrenci, danışmanın sürekli gözetimi altında tezini yazdı. Tez tamamlandıktan sonra danışman, tezin verilerinin yayımlanabilecek düzeyde ilginç olduğu sonucuna vardı. Öğrenci bu tezi bilimsel bir dergiye yollanacak şekilde yazamayacağı için, danışman yazdı.

A Öğrencinin bu durumda makalenin bir yazarı mı olması gerekir, yoksa yaptığı katkıların makalede bir dipnot olarak belirtilmesi mi gerekir? Sizin bu konuda karar vermenizi hangi faktörler etkiler?

B Eğer öğrencinin de yazar olması gerektiğine karar verirseniz, öğrencinin bu durumda ilk yazar mı yoksa ikinci yazar mı olması gerektiğini açıklayın.

KISIM İKİ

Betimsel Yöntemler

BÖLÜM DÖRT

Gözlem

BÖLÜMÜN ANA HATLARI

GENEL BAKIŞ
DAVRANIŞTAN ÖRNEK SEÇME
Zaman Örnekleme
Durum Örnekleme
GÖZLEME DAYALI YÖNTEMLER
DOĞRUDAN GÖZLEME DAYALI YÖNTEMLER
Müdahale Yapılmayan Gözlem
Müdahale Ederek Yapılan Gözlem
DOLAYLI (DİKKAT ÇEKMEYEN) GÖZLEME DAYALI YÖNTEMLER
Fiziksel İzler
Arşiv Kayıtları
DAVRANIŞIN KAYDEDİLMESİ
Kapsamlı Davranış Kayıtları
Seçilmiş Davranış Kayıtları
GÖZLEME DAYALI VERİLERİN ANALİZİ
Niteliksel Veri Analizi
Niceliksel Veri Analizi
GÖZLEME DAYALI ARAŞTIRMALAR ÜZERİNDE ELEŞTİREL DÜŞÜNME
Gözlemcinin Etkisi
Gözlemci Yanlılığı
ÖZET

GENEL BAKIŞ

Biz her zaman davranışları gözleriz. İtiraf edelim, çoğumuz insanları seyrederiz. Bu basitçe "dikizleme" düşkünü veya aşırı ölçüde meraklı olduğumuzdan değil, bazen insan davranışları kesinlikle ilginç olduğu içindir. İnsanların davranışları –jestleri, yüz ifadeleri, duruşları, kıyafet tercihleri– "beden dili" üzerine popüler kitaplarda vurgulandığı gibi, çok miktarda bilgi taşır (örneğin, Pease ve Pease, 2006). Sahip olduğumuz ve diğer insanlara sergilediğimiz şeyler (yatak odası duvarındaki posterler, raftaki kitaplar ve biblolar, dinlediğimiz müzik, sporla ilgili bir süs) bizim kişiliğimiz ve ilgilerimiz hakkında karar verirken bazen gözlemciye (farkında olmadan) ipuçları sağlar (Gosling, 2009). İster açık bir gülümseme veya isterse incelikli bir kur ritüeli olsun, diğer insanların davranışları, sıklıkla çabucak tanınan ipuçları verir. Doğrusunu söylemek gerekirse araştırmalar göstermiştir ki, pek çok yüz ifademiz bütün kültürlerde bilinen "evrensel" işaretlerdir (örneğin, Ekman, 1994). Bilim insanları da davranışlar hakkındaki pek çok şeyi öğrenmek için gözlemlerine dayanırlar [Yine de psikologların yeterince gerçek davranışları gözlemediklerine dair görüşler için Baumeister, Vohs ve Funder'e (2007) bakınız].

Bizim ve bilim insanlarının gündelik gözlemleri birbirinden çeşitli yönlerden farklıdır. Biz tesadüfî bir şekilde gözlem yaparız, gözlemimizi yanlı kılan etmenlerin farkında olmayabiliriz. Üstelik, gözlemlerimizi nadiren sistemli olarak kaydederiz. Onun yerine, kendi deneyimlerimiz (ve psikoloji araştırmaları) hafızamızın mükemmel olmadığını gösterdiği hâlde, yaşantılarımız hakkında kendi hafızamıza güveniriz.

Bilimsel gözlem, açık bir şekilde tanımlanmış koşullar altında, sistematik ve nesnel olarak ve dikkatli bir şekilde kayıt tutularak yapılır. Gözleme dayalı yöntemlerin birinci amacı davranışı betimlemektir. Bilim insanları davranışı tam olarak ve mümkün olduğunca doğru bir şekilde betimlemeye çalışırlar ama bu amaca ulaşmak için ciddi zorluklarla karşı karşıya kalırlar. Araştırmacılar için bir kişinin "bütün" davranışlarını gözlemek olanaksızdır. Dolayısıyla, bilim insanları, insanların davranışlarından örnekler gözlemlemeye dayanırlar. Ancak davranış sıklıkla ortaya çıktığı bağlam veya ortama bağlı olarak değişkenlik gösterir. Kendi davranışlarınızı düşünün. Okulda ve evde veya bir parti ile sınıfı karşılaştırdığınızda benzer şekillerde mi davranıyorsunuz? Diğer insanları –örneğin arkadaşlarınızı– gözlemek sizi, bağlamın önemli olduğu sonucuna götürdü mü? Bir çocuğun davranışlarının ebeveynlerden bazen birine bazen diğerine göre değiştiğini hiç gözlediniz mi? İşe yarar bir davranış tanımlaması, gözlemlerin pek çok farklı zaman ve pek çok farklı durumlarda yapılmasını gerektirir.

Bu bölümde, gözlemci-bilim insanlarının, bir davranış meydana gelirken, her zaman pasif şekilde o davranışın kaydını tutan kimseler olmadıklarını göreceksiniz. Bilim insanlarının gözlemleri için özel durumlar yaratarak müdahalede bulunmalarının nedenlerini açıklayacağız. Aynı zamanda, davranışın dolaylı olarak incelenmesini de ele alacağız. Bilim insanları fiziksel izler (örneğin, grafiti, ders kitapları üzerindeki altı çizilmiş yerler) ve arşiv kayıtları (örneğin, evlilik belgeleri, okul yıllıkları) üzerinde çalışarak davranış hakkında zengin bir anlayış kazanırlar. Gözleme dayalı verilerin kayıt ve analiz edilmesi için kullanılan yöntemlerden söz edeceğiz ve gözleme dayalı araştırma sonuçlarının yorumunu zorlaştıran hususları açıklayacağız. Gözlem, davranışlar

hakkında hipotezler kurmak için zengin bir kaynaktır; böylece gözlem, canlının neden öyle davrandığını keşfetme yolunda ilk aşama olabilir.

DAVRANIŞTAN ÖRNEK SEÇME

- Davranış eksiksiz olarak kaydedilmediği zaman, araştırmacılar davranışın temsil edici bir örneklemini elde etmeye çalışırlar.
- Gözlemlerin ne kadar genellenebileceği (dış geçerlik), davranıştan nasıl örnek seçildiğine bağlıdır.

Gözleme dayalı bir araştırma yürütmeden önce araştırmacılar, gözlemin nerede ve ne zaman yapılacağı hakkında bazı önemli kararlar almak zorundadır. Çoğu durumda yalnızca belirli bir ortamda, belirlenmiş koşullar altında ve belirlenmiş bir zamanda ortaya çıkan belli bir davranış gözlenebilir. Bir diğer ifadeyle, davranıştan örnek seçmek zorunludur. Bu örnek olası bütün davranışların geniş bir evrenini temsil eder. Araştırmacılar, gözlem yapmak için davranış evrenini temsil eden zaman, ortam ve koşulları seçerek bulgularını tüm evrene genelleyebilirler. Sonuçlar yalnızca gözlemin yapıldığı bu çalışmadaki katılımcı, zaman, ortam ve koşullara benzer olanlara genellenebilir. Temsil edici örneklemin temel özelliği onun, içerisinden seçildiği evren "gibi" olmasıdır. Örneğin, sene başındaki sınıf davranışlarının gözlemi, bu dönemdeki davranışları temsil edici olabilir. Fakat sene sonunda görülen tipik davranışlara ilişkin sonuçları vermeyebilir.

Temel Kavram

Dış Geçerlik araştırma sonuçlarının farklı evren, ortam ve durumlara genellenebilirliği demektir. Geçerliliğin doğruluk ile ilişkili olduğunu hatırlayın. Bir araştırmanın dış geçerliliğini belirlemek istediğimizde, araştırmanın bulgularının, bu araştırmada kullanılan katılımcı, ortam ve durumların dışındaki katılımcı, ortam ve durumlar için ne ölçüde kullanılabileceğini belirleriz. Şimdi, gözleme dayalı bulguların dış geçerliliğini arttırmak için zaman, olay ve ortam örnekleminin nasıl kullanılacağını açıklayacağız.

Zaman Örnekleme

- Zamandan örnek alma, araştırmacıların hem gözlem yapacakları zaman aralıklarını seçmelerini ve hem de bunun sistematik veya seçkisiz olup olmadığını ifade eder.
- Araştırmacılar nadiren görülen bir olay ile ilgilendiklerinde, davranışı örneklemek için olaydan örnek alırlar.

Temel Kavram

Araştırmacılar davranışın temsilî bir örneklemini belirlemek için genellikle zaman örnekleme ve ortam örneklemeyi birlikte kullanırlar. **Zaman örneklemede** araştırmacılar gözlemleri için çeşitli zaman aralıkları seçerek temsil edici bir örneklem oluşturmak isterler. Aralıklar sistematik olarak (örneğin, her haftanın ilk günü gözlem yapmak) veya seçkisiz veya her ikisi birden olacak şekilde seçilebilir. Çocukların sınıf içi davranışları gözlendiğinde zamandan nasıl örnek seçilebileceğini düşünelim. Eğer araştırmacılar gözlemlerini günün belirli bir zamanı (diyelim ki yalnızca sabah) ile sınırlamışlarsa, bulgularını okul gününün diğer zamanlarına genelleyemezler. Temsil edici örnek seçmeyi sağlamak için bir yaklaşım, gözlem aralıklarını sistematik olarak gün boyunca

KUTU 4.1

EHG İLE GÖZLEM YAPMAK: DİKKAT ÇEKMEYEN BİR ARKADAŞ

EHG bir "Elektronik olarak Harekete Geçen bir kayıt cihazıdır (electronically activated recording) ve araştırmacıların katılımcıların gündelik hayatına işitsel bir pencere açmaktadır. Araştırmacılar, katılımcıların doğal ortamlarındaki çeşitli psikolojik olayları dinleyerek, davranış ile bireyin dil kullanımı hakkında zengin bilgiler elde ederler. Bu yöntem, katılımcıların kendi davranışları hakkındaki öz bildirimlerini (self-report) tamamlayıcı niteliktedir; öz bildirim, katılımcının kendi öznel yorumlarından etkilenebilir ve ayrıca, katılımcı kendi davranışlarının tam olarak farkında olmayabilir. (Mehl ve Pennebaker, 2003). Araştırmacılar EHG kullanarak:

- Kâğıt-kalem testleriyle ölçülen kişilik ve mutluluk düzeyi bildirimlerini, bireylerin iletişim tarzlarıyla ilişkilendirdiler (Mehl, Vazire, Hollearn ve Clark, 2010).
- "Kadınlar erkeklerden daha mı konuşkandır ?" sorusunu cevapladılar (Mehl ve ark. 2007).
- Üniversite öğrencilerinin "toplumsal dünyası"nın istikrarlılığını incelediler (Mehl ve Pennebaker, 2003).

EHG aynı zamanda insanların doğal ortamlarındaki fiziksel sağlıkları ve psikolojik iyi oluş hâllerini değerlendirmek için kullanışlı bir strateji sağlayabilir. Sağlıkla ilgili ifadelerin (örneğin, duygu ifadelerinin) ve davranışların işitsel kayıtları, geleneksel sağlık değerlendirmelerinin eksiklerini giderebilir (Mehl, Robbins ve Deters, 2012). EHG, araştırmacılara doğal davranışın önemli boyutlarını kulaklarıyla dinleyebilmek için benzersiz bir fırsat sağlar. Yöntem, sonuçların dış geçerliliğini güçlendirmek için hem zaman ve hem de ortamdan örnek alarak desteklenir.

programlamaktır. Gözlemler, 4 kez, 8:30, 10:30, 12:30 ve 14 30'da başlayacak şekilde, 30 dakikalık sürelerle yapılır. Benzer bir durumda kullanılacak seçkisiz bir zaman tekniği, 30 dakikalık dört zaman dilimini, seçkisiz olarak gün içine dağıtmaktır. Farklı bir seçkisiz tarife, her gün gözlem yapmaktır.

Zamanı günden güne değişkenlik gösterebilir fakat uzun bir süre devam edildiğinde davranış örnekleri bir okul gününün bütün zamanlarından alınmış olur. Son 20 yılda, davranış örneği seçilirken elektronik cihazların kullanımı artış göstermiştir. Örneğin elektronik çağrı cihazları, seçkisiz bir zaman tarifesine göre gözlemciye sinyal göndermeye programlanabilmektedir (normal uyku zamanları hariç). Orta sınıftan gençlerle yapılan bir çalışmada araştırmacılar, ergenlerin günlük hayatlarından "16.477 seçkisiz an"daki deneyimlerine ilişkin bildirimler aldılar (Larson ve ark., 1996). Bu "deneyim örnekleme yöntemi" ile elde edilen sonuçlardan yola çıkarak Larson (2000; 2011), gençlerin olumlu gelişimine katkıda bulunabilecek hedefleri belirlemek için, gençlerin yeterliliklerini geliştiren öğrenme çabalarını araştırdı (örneğin, gönüllülük, spor, sanat).

Bir diğer ilginç yöntem, bireyin gündelik etkinliklerine işitsel bir giriş sağlayan EHG kullanımıdır (Kutu 4.1'e bakınız). Katılımcılar 12.5 dakika aralıklarla ve 30 sn. boyunca çevrelerindeki sesleri kaydetmeye programlanmış küçük bir ses kayıt cihazı kullanırlar. Jüri bu kayıtları dinler ve her bir aralığın sonunda, yer (ev, lokanta vb.), etkinlik (müzik dinleme, bilgisayarda çalışma vb.), etkileşim (yalnız, diğerleriyle konuşma) ve ruh durumu (gülme, ağlama vb.) ölçütlerine göre katılımcının doğal ortamındaki sosyal çevreyi kodlar. Ayrıca, EHG ile tespit edilen konuşma dökümleri, bilgisayar destekli dil analizine tâbi tutulur (Mehl ve Holleran, 2007'ye bakınız).

Sistematik ve seçkisiz zaman-örnekleme işlemleri, çoğunlukla birlikte kullanılır. Gözlem aralıkları sistematik olarak düzenlenir fakat gözlem, seçkisiz aralıklarla

yapılır. Örneğin elektronik çağrı cihazı, her üç saatte bir (sistematik) fakat seçkisiz belirlenmiş bir süre sinyal vermeye/almaya programlanabilir. Hangi zaman-örnekleme işlemi kullanılırsa kullanılsın, zamandan örnek almanın amacı canlının olağan davranışlarını temsil edecek bir davranış seçkisi sağlamaktır.

Zamandan örnek alma, nadiren ortaya çıkan olaylarla ilgilenildiği zaman, davranışı örneklemek için etkili bir yöntem değildir. Az rastlanan olayları incelerken zaman örneklemeyi kullanan araştırmacılar, bu olayı bütünüyle kaçırabilirler. Veya eğer olay çok uzun sürüyorsa, zaman örnekleme yöntemi araştırmacının olayın başı veya sonu gibi önemli kısımlarını kaçırmasına neden olabilir. *Olay* örnekleme'de gözlemci, önceden belirlenmiş bir olayla her karşılaştığında onu kaydeder. Örneğin, bir hayvanın yemek yeme davranışını incelemek isteyen bir araştırmacı, olay örnekleme yöntemini kullanabilir. Spor psikologları, atletik yarışmalarda yalnızca belirli davranışları kaydedebilir (örneğin, Bowker ve ark., 2009). Gözlemin ne zaman yapılacağını, bu belirli olay belirler.

Olay örnekleme aynı zamanda, doğal veya diğer felaketler gibi beklenmedik şekilde meydana gelen olaylar gözleneceği zaman da kullanışlıdır. Mümkün olduğunca, gözlemciler ilgi çekici bir olay olduğunda veya olma ihtimali olduğunda orada olmaya çalışırlar. Olay örnekleme, beklenmedik veya seyrek gözlenen olaylar için etkili ve yeterli olmakla birlikte, olay örneklemenin kullanımı davranışın kayıt edilmesinde kolayca yanlılığa neden olabilir. Örneğin, olay örnekleme gözlemcinin sadece "çok uygun" veya sadece olayın mutlaka ortaya çıktığı zamanlardan örnek seçmesine yol açabilir. Davranıştan sadece bu zamanlarda örnek alınmasının bir sonucu olarak, diğer zamanlardaki benzer davranışları temsil etmeyebilir. Temsil edici bir örnek sağlamak üzere yine de bir başka örnekleme işlemi var: Durumdan örnek alma.

Durum Örnekleme

- Durum örnekleme, farklı yer ve farklı hâl veya koşullardaki davranışların incelenmesini kapsar.
- Durum örnekleme, bulguların dış geçerliliğini arttırır.
- Bir durumun içinde, yalnızca ortamdaki bazı bireyleri gözlemek için katılımcı örnekleme de kullanılabilir.

Temel Kavram

Araştırmacılar durumdan örnek alarak, gözleme dayalı bulguların dış geçerliliğini büyük ölçüde yükseltebilirler. **Durum örnekleme** mümkün olan en fazla farklı koşul ve hâlde ve farklı yerlerde davranışın gözlenmesini içerir. Çeşitli durumlardan örnek alarak, araştırmacılar sonuçlarının yalnızca belirli hâl veya koşullara mahsus olma olasılığını azaltırlar. Örneğin, hayvanlar, hayvanat bahçesindeyken vahşi hayatta veya farklı yerlerde olduğuna benzer şekilde davranmazlar. Anne ve bebek şempanzeler arasındaki bakışmaların çalışıldığını düşünelim. Ortak göz hareketleri, insanlarda olduğu gibi maymunlarda da görülür fakat şempanzeler üzerine bir çalışmada bu davranışın sıklığı Japonya ve Amerika'da gözlenen hayvanlar arasında farklılaşmıştır (Bard ve ark., 2005). Benzer şekilde, insanlar arasındaki bakışların kültürden kültüre değişkenlik gösterdiği bilinmektedir (örneğin, McCarthy, Lee, Itakura ve Muir, 2006).

Bir davranışın daha etkili bir şekilde gözlenebileceği pek çok durum vardır. Örneğin, eğer araştırmacılar yemek salonunun en kalabalık saatlerinde öğrencilerin yemek tercihlerini gözlüyorlarsa, öğrencilerin tamamını gözleme imkanları

olamaz. Bu örnekte, diğerlerinde olduğu gibi, araştırmacılar hangi öğrencileri gözleyeceklerini belirleyerek katılımcı örnekleme yaparlar. Zaman örnekleme prosedürüne benzer şekilde, araştırmacılar öğrencileri sistemli şekilde (örneğin, her 10'uncu öğrenciyi) seçebilirler veya seçkisiz olarak seçebilirler. Muhtemelen artık aşina olduğunuz gibi, katılımcı seçmenin amacı, örneklemin evreni temsil etmesini sağlamaktır; yani, bu örnekte, yemek salonunda yemek yiyen bütün öğrencileri.

GÖZLEME DAYALI YÖNTEMLER

- Gözleme dayalı yöntemler, doğrudan gözlem ve dolaylı gözlem olarak sınıflanabilir.

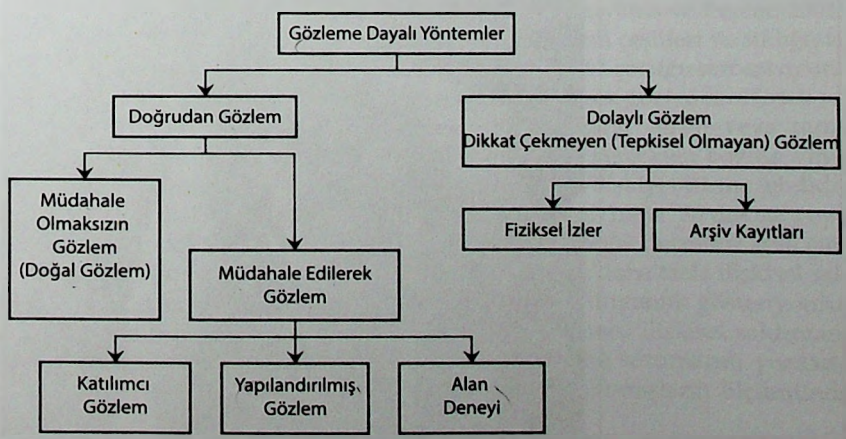
Araştırmacılar çoğu kez davranışı oluştuğu esnada gözlerler; buna doğrudan gözlem denilir. Bununla beraber gözlemler dolaylı da yapılabilir; örneğin, araştırmacılar fiziksel izler veya arşiv kayıtları kullanarak geçmiş davranışlara ilişkin kanıtları inceledikleri zaman olduğu gibi. Bu dolaylı (veya dikkat çekmeyen) gözlemdir. Şekil 4.1'de gözleme dayalı yöntemler gösterilmiştir. İlk önce doğrudan gözlem yöntemlerini, daha sonra dolaylı (dikkat çekmeyen) gözlem yöntemlerini ele alacağız.

DOĞRUDAN GÖZLEME DAYALI YÖNTEMLER

- Doğrudan gözleme dayalı yöntemler, "müdahale yapılmayan gözlemler" ve "müdahale yapılan gözlemler " olarak sınıflanabilir.

Davranış doğrudan gözlendiği zaman, araştırmacılar gözledikleri duruma ne ölçüde müdahale edebileceklerine ilişkin bir karar verirler. Demek ki müdahale, araştırmacıların gözlem için bir bağlam yaratma veya onu değiştirme çabasını kastetmektedir. Müdahalenin sınırları, sıfır müdahaleden (müdahalesiz gözlem) doğal ortam deneyi yürütülmesine kadar uzanan geniş bir değişkenlik gösterir.

ŞEKİL 4.1 Gözleme Dayalı Yöntemlerin Akış Şeması.



KUTU 4.2

GÖZLEM: BİR DİĞER BAKIŞ

Doğal ortamlardaki davranışları gözleyen tek araştırmacılar psikologlar değildir. Gözlem, biyolojinin bir dalı olan etolojinin temel yöntemidir (Eibl-Eibesfeldt, 1975). Etologlar, canlıların doğal ortamlarıyla ilişkili davranışlarını incelerler. Tipik olarak hayvanları doğal ortamlarında uzun süreler gözleyerek kaydederler. İnsan davranışlarının belirlenmesinde doğuştan gelen mekanizmaların rolü hakkındaki spekülasyonlar, etologlar arasında da yaygındır.

Hayvanlarda çiftleşme ve üremeyi araştıran araştırmacılar, aynı cinsten hayvanların cinsel davranışları ve ebeveynlik davranışlarını ortaya koyan kanıtları yorumlamakta güçlük çekiyorlar (Mooallem, 2010). Pek çok biyoloğun insan cinselliğiyle karşılaştırmaktan kaçınmasına rağmen, hemcinslerin cinsel davranışları ve ebeveynlik davranışları gözlemleri bir hayli tartışmaya yol açmıştır (Şekil 2'ye bakınız). Eşcinsellikle ilişkilendirilen bu sosyo-politik tartışmanın her iki tarafındakiler, hemcins hayvanlar arasında cinsel davranışların varlığına ilişkin kanıtları, kendi gündemlerini desteklemek için kullandılar. Ancak, bilimsel gözlemin ayırıcı özelliği, nesnel ve yanlılıktan arınmış olmaktır –politik gündemler dâhil. Ama yine de hayvan davranışlarını kendi bağlamında, kendi maksadına uygun olarak yorumlamak yerine, hayvan cinselliğini eşcinsellik veya lezbiyenlik gibi insanlara özgü terimleri kullanarak yorumlayanlar vardır.

Aynı cinsten canlıların cinsel davranışlarını anlama konusundaki sorunu, evrimsel biyolojinin kalbine uzanır; şöyle ki, bütün evrimsel-uyum davranışları bir temel amaca yöneliktir: Genlerini aktarmak. Bununla birlikte, biyologlar yakın zamanlarda hemcins hayvanlar arasındaki ebeveynlik ve cinsel davranışların, uyumun bir ürünü olabileceğini öne süren bir teori geliştirdiler. Nesnel gözlem ve teori kurma süreci, bütün bilimler için temel bir süreçtir. Yine de Bölüm 1'de işaret ettiğimiz gibi, bilim kültürel bir bağlamda meydana gelir ve bu durum, bazı insanların bu sürecin sonuçlarını yorumlarken nesnellikten uzaklaşmalarına yol açabilir.

ŞEKİL 4.2 Ve Tango Üçledi (And Tango Makes Three, Richardson ve Parnell, 2005) isimli çocuk kitabı Central Park Hayvanat bahçesinde gözlenen iki erkek penguenin bir bebek pengueni yetiştirmesi hikayesine dayalı olarak yazılmıştır. Amerikan Kütüphaneler Birliğinin raporlarına göre, kitap 2009'da en sık yasaklanan kitap olmuştur.



Müdahale Yapılmayan Gözlem

- Doğal gözlemin amacı, davranışı olduğu hâliyle tarif etmek ve değişkenler arasındaki ilişkileri incelemektir.
- Doğal hâlde gözlem, laboratuvar bulgularının dış geçerliliğini sağlamaya yardım eder.
- Deneysel kontrolü engelleyen ahlâkî ve etik hususlar varsa, doğal gözlem önemli bir araştırma stratejisidir.
- İnternetteki davranış, müdahale edilmeden gözlenebilir.

Temel Kavram

Gözlemci tarafından herhangi bir müdahale *yapılmadan*, davranışın doğal ortamında doğrudan gözlenmesine, **doğal gözlem** denilir. Bir gözlemci bu gözlem yöntemini kullandığında, doğal olarak oluşmuş bir olayın pasif bir kaydedicisi gibi davranır. Doğal bir ortamda oluşan sıradan bir davranışı düşünelim; bu davranış, bir psikoloji laboratuvarındaki gibi gözlenmek amacıyla özel olarak hazırlanmamıştır. Örneğin, Matsumoto ve Willingham (2006) atletleri Olimpik judo yarışmasının doğal ortamında (bu atletler için) gözlediler. Kutu 4.2 etoloji alanında doğal gözleme dayalı son bulguları anlatıyor.

Doğal ortamlarda gözlem, diğer işlevlerinin yanı sıra, laboratuvarda elde edilen bulguların dış geçerliğini sağlamanın bir yoludur; laboratuvarı "gerçek dünyaya" taşıma işlevi görür. Örneğin, soru formları ve görüşmelerden elde edilen bulgular gösteriyor ki ebeveynin depresyonu ile olumsuz çocukluk yaşantıları arasında bir ilişki vardır. Araştırmacılar bu ilişkinin, çocukların günlük hayatlarını nasıl etkileyeceğini açıklamak için 35 çocuğun doğal gözleminden yararlandılar (Slatcher ve Trentacosta, 2012). Davranış ve konuşmaları izlemek için bir yıl arayla iki kez 1-gün süreyle EHG sisteminin çocuk versiyonunu kullandılar (Kutu 4.1.'e bakınız). Sonuçlar, daha yüksek düzeyde ebeveyn depresyon belirtilerinin (soru formundan gelen veriye göre) çocuklardaki daha yüksek düzeyde problemli davranışla (ağlama, öfkeli davranışlar) bağlantılı olduğunu gösterdi.

Doğal ortamda gözlemin esas amacı, davranışı olduğu gibi betimlemek ve mevcut diğer değişkenlerle ilişkilerini araştırmaktır. Gelişim psikologları, sınıf ve oyun alanlarındaki doğal ortamlarda, çocukluk dönemindeki zorba davranışları da içeren saldırganlığın çeşitlerini ve sıklığını araştırdılar (örneğin, Crick ve ark. 2006; Drabick ve Baugh, 2010; Hartup, 1974; Hawkins ve Pepler, 2001; Ostrov ve Keating, 2004). Özellikle akran saldırganlığının çeşitleri ve sıklığıyla ilgilendiler. Bir araştırmada araştırmacılar 18 ay ve üzeri 91 çocuğu serbest oyunları esnasında fiziksel ve ilişkisel saldırganlığı kaydederek gözlediler (Crick ve ark., 2006). Fiziksel saldırganlık bir diğerine fiziksel zarar verme veya zarar verme tehdidi olarak, ilişkisel saldırganlık ise ilişkiyi gerçekten bozma veya bozmakla tehdit etme (sosyal dışlama, sessiz davranma, ilişkiyi bitirme tehdidi) olarak tanımlandı. Her bir kız ve erkek çocuk, dört farklı hafta, 80 dakika veya toplamda 320 dakika gözlendi. Beklendiği gibi, saldırganlık öncelikle aynı cinsiyetten arkadaşa yöneltiliyordu; kızlar erkeklere oranla daha fazla ilişkisel saldırganlık, erkekler ise kızlardan daha fazla fiziksel saldırganlık gösteriyordu. Daha da önemlisi, araştırmanın yapıldığı dönem boyunca ilişkisel saldırganlık oldukça istikrarlı kaldı ve ilerideki akran reddetme sorunlarını yordadı. Araştırmacılar gözleme dayalı yöntemlerin, sosyal davranışların ölçümünde

diğer değerlendirme türlerine (akran veya öğretmen bildirimleri gibi) oranla daha nesnel olduğunu savunmaktadırlar.

Çocuk saldırganlığı araştırması, niçin araştırmacıların çalışılan davranışa deneysel müdahale yerine doğal gözlem kullanmayı seçtiklerini gösteriyor. İnsan davranışlarının deneysel kontrol sağlamayı engelleyen ahlâkî veya etik yönleri vardır. Örneğin, araştırmacılar, erken çocuklukta tecrit edilmenin sonraki duygusal ve psikolojik gelişimle ilişkisiyle ilgileniyor olsunlar. Ancak, eğer çocukları ailelerinden almak suretiyle onları tecrit etmeye teşebbüs ederlerse onlara ciddi şekilde karşı çıkarız. Eğer çocukluktaki tecrit incelenecekse, veri toplamak için alternatif bir yöntem düşünülmek zorundadır. Örneğin, erken tecrit edilmenin sonraki gelişim üzerindeki etkileri hayvan deneklerle deneysel olarak çalışılmıştır (Harlow ve Harlow, 1966); vahşi çocuklar olarak isimlendirilen insan kültürünün dışında muhtemelen hayvanlarla büyümüş çocuklar gözlenmiştir (Candland, 1993); ebeveynleri tarafından alışılmadık tecrit koşullarına maruz bırakılan çocuk vakaları incelenmiştir (Curtiss, 1977) ; kurum bakımı altındaki çocukların doğrudan ve sistematik gözlemleri yapılmıştır (Spitz, 1965). Ahlâkî ve etik yaptırımlar, çocuklarda saldırganlığın niteliğini incelerken de uygulanır. Çocukların, sadece davranışlarını kaydetmek amacıyla kasıtlı olarak taciz edildiklerini görmek istemeyiz. Bununla birlikte çocukları gözleyen herkes bilir ki kendiliğinden oluşan çok sayıda saldırganlık vardır. Doğal gözlem, ahlâkî ve etik sınırlar dâhilinde kalmak koşuluyla, çocukların saldırganlıkları hakkında bilgi edinmek için kullanışlı bir yöntem olabilir.

Araştırmacılar, doğal gözlem için interneti de kullanırlar. İnternet toplulukları, sohbet odaları, haber grupları ve aynı anda oynanan çok oyunculu gruplar (örneğin, MUD veya MMORG) gibi internet toplulukları, çok çeşitli davranışları araştırmak için doğal ortamlar sunar (örneğin, Whitlock, Powers ve Eckenrode, 2006). Bu ortamlar özellikle niteliksel analizler için uygun görünmektedir (Bölüm 2'ye ve bu bölümün sonundaki niteliksel analizler hakkındaki tartışmaya bakınız). Bir ergen sohbet odasında pasif gözlem, örneğin "kimlik ve cinsiyet meselesinin nasıl inşa edildiğini araştırmak" için kullanılmıştır (Subrahmanyam, Greenfield ve Tynes, 2004, s. 651). Böyle çalışmalar, gizliliğin korunamadığı durumlarda katılımcılara verilecek potansiyel zararlar veya gözlemcinin varlığı fark edildiğinde topluluğa verilecek zararlar üzerinde dikkatlice düşündükten sonra yapılmalıdır. Kullanıcıların bu alanı (interneti) "umumî" mi yoksa "özel" mi algıladıkları hususu, özellikle önemlidir (bk. Eysenbach ve Till, 2001).

Müdahale Ederek Yapılan Gözlem

- Çok sayıda psikolojik araştırma müdahale edilerek yapılan gözlemi kullanır.
- Müdahaleli gözlemin üç şekli; katılarak gözlem, yapılandırılmış gözlem ve alan deneyleridir.
- İster "gizlenmiş", isterse "gizlenmemiş" olsun, katılımcı gözlem, araştırmacıların bilimsel gözleme genellikle açık olmayan davranış ve ortamları gözlemesine izin verir.
- Eğer bireyler gözlendiklerini bildikleri zaman davranışlarını değiştiriyorlarsa (tepkisellik), gözlenen davranışlar artık onların normal davranışlarını temsil etmez.

- Klinik ve gelişim psikologları tarafından sık kullanılan yapılandırılmış gözlem, doğal gözlem kullanılarak gözlenmesi zor olabilecek davranışları kaydeden bir yöntemdir.
- Alan deneyinde araştırmacılar, bir veya daha fazla bağımsız değişkeni, doğal ortamdaki davranışa olan etkisini belirlemek için değişimler.

Bilim insanlarının doğayı kurcalamayı sevdikleri, bir sır değildir. Etkilerini gözlemek veya belki de bir teoriyi sınamak için müdahalede bulunmaktan hoşlanırlar. Müdahale edilen psikolojik araştırmalar, edilmeyenlerden çok daha fazladır. Araştırmacılar doğal ortamlara müdahale edecekleri zaman, kullanmayı tercih edecekleri üç önemli gözlem yöntemi vardır: Katılımcı gözlem, yapılandırılmış gözlem ve alan deneyi. Bu üç yöntemde müdahalenin doğası ve derecesi farklıdır. Şimdi her birisini sırasıyla ele alacağız.

Temel Kavram

Katılımcı Gözlem **Katılımcı gözlemde** gözlemciler ikili rol oynar: İnsanların davranışlarını gözlerler ve gözlem yaptıkları ortama etkin olarak katılırlar. Gizlenmemiş katılımcı gözlemde, gözlenen bireyler, gözlemcinin davranışları hakkında bilgi toplamak amacıyla orada bulunduğunu bilirler. Bu yöntem sıklıkla grup üyeleri ile çalışarak ve yaşayarak grubun davranışları ve kültürünü anlamak isteyen antropologlar tarafından kullanılır.

Gizlenmiş katılımcı gözlemde, gözlenen, gözlendiğini bilmez. Anlaşılacağı üzere, insanlar davranışlarının kaydedildiğini bildikleri zaman, her zaman davrandıkları şekilde davranmazlar. Daha sonra tartışacağımız gibi, davranışın gözlenmesinde en temel sorun tepkiselliktir. **Tepkisellik**, insanların gözleniyor olduklarını bilmelerine, normal davranışlarını değiştirerek tepki vermelerine denir. O yüzden araştırmacılar, gözlendiklerini anladıkları zaman insanların davranışlarını değiştireceklerini düşünüyorlarsa, gözlemci olarak rollerini gizlemeye karar verebilirler. Gizlenmiş katılımcı gözlem, etik sorunlara neden olur (örneğin, mahremiyet ve bilgilendirilmiş onam gibi) bu nedenle bu sorunlar, araştırma başlamadan önce ele alınmalıdır. Bu etik sorunları Bölümde 3'te ele almıştık ve bu bölümün sonlarında da tartışacağız.

Temel Kavram

Katılımcı gözlem, genellikle bilimsel gözleme açık olmayan ortamlar için gözlemciye bir fırsat sağlar. Zihinsel rahatsızlıkların psikiyatrik teşhisi ve hastane tedavisi ile ilgili klasik çalışmasında Rosenhan (1973), hastane yönetiminin bilgisi dahilinde, gizlenmiş katılımcı gözlemcilerle çalıştı. Muayene esnasında katılımcıların hepsi de bazı sesler işitmek gibi genel bazı şikâyetler bildirdiler. Bu sahte hastaların bir çoğu şizofreni teşhisi aldı. Hastaneye yatırıldıktan hemen sonra, katılımcı gözlemci hastalar bu sahte belirtileri göstermeye son verdiler ve "aklı başında" insanlar olarak, hastaneden bırakılmalarının ne kadar zaman alacağını görmeyi beklediler. Hastaneye yatmalarıyla beraber gözlemlerini de kaydettiler. Araştırmacılar bir hafta ile 52 gün arasında hastanede kaldılar. Taburcu edildiklerinde kendilerine, şizofrenilerinin "hafiflemiş" olduğu söylendi. Açık bir şekilde bu sahte hastalar, bir kere şizofreni ile etiketlendikten sonra, bu etiket onlara yapıştı kaldı.

Ancak Rosenhan'ın (1973) bu sonuçlarına karşı çıkmak için nedenler vardır (örneğin, Spitzer, 1976; Weiner, 1975). Örneğin, psikiyatristler "hafifledi" ifadesiyle, basitçe bireyin belirtileri artık göstermediğini kastetmiş olabilirler.

Şizofren bir bireyin belirtilerini tekrarlaması beklendiğinden "hafiflemiş" teşhisi, akıl sağlığı uzmanlarına bireyin davranışlarını anlamak için daha çok çaba sarf etmeleri gerektiğini söylemektedir. Aynı zamanda Rosenhan'ın çalışması, akıl sağlığı uzmanlarına, nasıl teşhis koyduklarına, kişisel ve teorik yanlılıkların nasıl rol oynadığını da gözeterek, daha dikkatli olmaları hususunda yol göstermektedir.

Gizlenmiş katılımcı gözlemciler, çalışmaya dâhil olan insanlarla benzer deneyimlere sahip oldukları için, bireyler veya grup hakkında derin bir anlayış ve iç görü kazanırlar. Rosenhan'ın çalışmasındaki sahte hastalar, mesela şizofreni tanısı almanın nasıl bir şey olduğunu ve topluma yeniden dâhil olmanın ne kadar uzun sürebileceğini anladılar. Rosenhan'nın (1973) çalışmasının önemli bir diğer katkısı, kurum ortamlarında oluşabilen insanî tutumdan uzaklaşmanın bir resmini ortaya koymasıdır.

Bazen katılımcı gözlemcinin rolü, başarıyla yürütülen bir çalışmada ciddi sorunlara neden olabilir. Mesela gözlemciler eğer gözledikleri ortam ve insanlarla çok yakınlaşırlarsa bilimsel nesnelliklerini kaybedebilirler. Örneğin bir kriminoloji uzmanı, gizlenmemiş katılımcı gözlemci olarak polis akademisi eğitimini incelemeye gitti ve yoğun suç bölgelerine üniformalı devriye memuru olarak gönderildi (Kirkham, 1975). Memur olarak geçirdiği bu deneyim, tutumlarında, kişiliğinde, duygu durumunda ve davranışlarında beklenmedik ve çarpıcı değişimlere yol açtı. Bir araştırmacı olarak kendisinin, "cezaya düşkünlük, yaygın bir alaycılık ve diğerlerine güvensizlik, kronik sinirlilik ve açığa çıkmış düşmanlık, ırkçılık ve suçlular ve suç tehdidini aşacak kadar yaygın kişisel kaygı" gösterdiğine dikkat çekti (s. 19). Böyle bir ortamda katılımcı gözlemci, ortama kendisinin katılmasının, özellikle bağlılığındaki artış nedeniyle, objektif bir rapor oluşturmaya bir tehdit olduğunun farkında olmak zorundadır.

Katılımcı gözlemin bir diğer potansiyel sorunu, gözlemcinin araştırmaya dâhil olan insanların davranışlarını etkileyebilme olasılığıdır. Muhtemeldir ki katılımcı gözlemci insanlarla etkileşime geçecektir, kararlar alacak, etkinlikler başlatacak, sorumluluk üstlenecektir. Bir diğer ifadeyle ortamdaki herkes gibi davranacaktır. Durumda ve katılımcılarda gözlenen değişimler bu katılımdan etkilenmiş olabilir mi? Eğer insanlar katılımcı gözlemciden dolayı normal olarak davrandıkları gibi davranmadılarsa, bu sonuçları diğer ortamlara genellemek zordur.

Katılımcı gözlemcinin gözlenen davranışlar üzerindeki etkilerinin boyutları kolayca değerlendirilemez. Çok sayıda faktör göz önüne alınmalıdır; katılım gizlenmiş mi gizlenmemiş mi olmalıdır? Grup büyüklüğü ne olmalıdır? Gözlemcinin gruptaki rolü ne olmalıdır? Gözlem altındaki grup küçük veya katılımcı gözlemcinin etkinlikleri göze çarpıcıysa, gözlemcinin insanların davranışlarına anlamlı etkisi çok daha fazla olacaktır. Bu yüzden katılımcı gözlem, bilimsel araştırmaya çok açık olmayan ortamlara girme fırsatı sağlamasına rağmen, bu tekniği kullanan bir gözlemci muhtemel nesnellik kaybı ve bir gözlemcinin davranış üzerindeki olası etkileri ile baş etmenin yollarını aramak zorundadır.

Katılımcı gözlem, tipik olarak dışarıya açık olmayan çeşitli internet topluluklarına girmek üzere kullanılmaktadır. Örneğin, bir araştırmacı, "Beyaz Irkçılar internet sohbet odası"na "meraklı çırak" olarak girdi ve Afrika kökenli

Amerikalara yönelik nefret suçlarının analizini yaptı (Glaser, Dixit ve Green, 2002). Bir diğer örnekte araştırmacı, bir çeşit psikolojik destek hizmeti veren "pro-ana" internet topluluğuna, pro-ana sitesi üzerinden gizlenmiş katılımcı gözlemci olarak katıldı (Brotsky ve Giles, 2007). (Pro-ana, anoreksiya nevroza yeme bozukluğunun kısaltılmışıdır).

İnternette katılımcı gözlem, bu yöntemleri tanıtırken sözünü ettiğimiz, gizlenmiş gözlemci mevcutken katılımcının bilgilendirilmiş onam vermesinin mümkün olmaması gibi etik sorunlar ortaya çıkarır. Bu yöntem, genellikle gözlenmeyi kabul etmeyen bireyleri gözleme ve laboratuvarda test edilebilecek hipotezler geliştirme olanağı sağlamakla birlikte, hem gözlem ve hem de verinin kodlanması için uzun zaman harcamayı gerektirir. Ayrıca araştırmacılar her zaman sonuçlarını diğer sanal topluluklar veya yüz-yüze etkileşim gözlemlerine genelleyemeyebilir (Online katılımcı gözlemin kullanımı ile ilgili yararlı öneriler ve bu tür mevzulara ilişkin tartışmalar için Utz, 2010'a bakınız).

Yapılandırılmış Gözlem Kolayca sınıflandırılması mümkün olmayan ve bir müdahale içeren çok sayıda gözleme dayalı yöntem vardır. Bu işlemler doğal hâlde gözlemden farklıdır, çünkü araştırmacılar gözledikleri durum üzerinde bazı kontroller uygulayarak müdahalede bulunurlar. Durum üzerindeki kontrol ve müdahalenin derecesi, alan deneylerindekinden azdır (Bu bölümde kısaca, Bölüm 6'da ise daha ayrıntılı olarak tartışacağız.). Bu prosedürü **yapılandırılmış gözlem** olarak isimlendireceğiz. Sıklıkla gözlemciler bir olayın ortaya çıkmasına veya bir durum oluşmasına neden olarak bir müdahalede bulunurlar ve böylece olayları kolayca kayıt edebilirler.

Temel Kavram

Araştırmacılar belirli bir davranışı bütünüyle inceleyebilmek için ayrıntılı bir dizi işlem tasarlayabilir. 'Dikkat körlüğü olarak adlandırılan olguyla ilgili bir çalışmada, araştırmacılar insanların cep telefonu kullanırken alışılmadık durumları yakalama becerileri üzerine çalıştılar (Hyman, Boss, Wise, McKenzie ve Caggiano, 2009). Dikkat körlüğü, insanların özellikle dikkatleri herhangi bir yere odaklanmışken –cep telefonu gibi– çevrelerindeki yeni ve ayırt edilebilir bir uyarımı yakalama başarısızlığı gösterdiklerinde ortaya çıkar. Bu çalışmada, araştırmacılar bir *işbirlikçi* (confederate) kullandılar. Bu yöntemde, araştırma ortamındaki birisiyle, gözlenecek davranışı yaratmak üzere belirli bir şekilde davranması için önceden anlaşma yapılır. Hyman ve arkadaşlarının çalışmasında, bir işbirlikçi palyaço kıyafeti giyerek üniversite kampüs meydanındaki büyük bir heykelin çevresinde tek tekerlekli bir bisiklet sürdü (Şekil 4.3'e bakınız). Bir saatlik bir süre boyunca palyaço, bu meydanda bisiklet sürdü. Araştırmacılar, meydanda yürüyen yayalara alışılmadık bir şey görüp görmediklerini sordular. Eğer "evet" dedilerse, ne görmüş olduklarını sordular. Eğer yayalar palyaçodan söz etmedilerse, özellikle tek tekerlekli bir bisiklete binen bir palyaço görüp görmediklerini sordular.

Bu yapılandırılmış gözlem işlemi, cep telefonu kullanma esnasında, dikkat körlüğünün daha çok gösterilip gösterilmeyeceğini belirlemek için tasarlanmıştı. Araştırmacılar yayaları dört gruba böldüler; cep telefonu kullananlar, sadece yürüyenler (elektronik araçsız), sadece müzik dinleyerek yürüyenler. (örneğin, MP3 ile) veya çift olarak yürüyenler. Sonuçlar gösterdi ki, cep telefonu kullananlar palyaçoyu fark etme ihtimali en düşük olanlardı. Cep telefonu

ŞEKİL 4.3 Hyman ve arkadaşlarının (2009) dikkat körlüğü çalışmasındaki tek tekerlekli bisikletteki palyaçonun fotoğrafı



kullananların yalnızca %25 'i palyaçoyu fark etti; bu oran çift olarak yürüyenlerde %71, müzik dinleyenlerde %61 ve tek başına yürüyenlerde ise %61 oldu. Müzik dinleyenleri ya da bir başkasıyla yürüyenlerin palyaçoyu fark etme ihtimallerinin yüksekliğine dikkat edin. Bu, cep telefonu kullanıldığı zaman bölünen dikkatle bağlantılı olan belirli bir şeylerin, dikkat körlüğü ile ilişkili olabileceğini düşündürmektedir. Hyman ve arkadaşları (2009) eğer böyle yüksek düzeyde dikkat körlüğü basit bir etkinlik olan yürüme esnasında bile var ise, araba sürerken cep telefonu kullanarak ortaya çıkan körlüğün ne kadar büyük olabileceğine dikkat çektiler.

Yapılandırılmış gözlem, Hyman ve arkadaşlarının (2009) araştırmasında olduğu gibi, doğal ortamlarda veya laboratuvar koşullarında düzenlenebilir. Klinik psikologlar, ebeveyn-çocuk etkileşiminin davranışsal değerlendirmesini yapacakları zaman sıklıkla yapılandırılmış gözlem kullanırlar. Örneğin, araştırmacılar kötü muamele (örneğin, taciz, ihmal) görmüş ailelerden gelen ve gelmeyen anne ve çocukların birlikte oynadıkları oyunları gözlediler (Valentino, Ciccethi, Toth ve Rogosh, 2006). Anneler, laboratuvar ortamında araştırmacıların düzenlediği farklı bağlamlarda çocukları ile etkileşim hâlindeyken, tek yönlü bir aynadan izlenerek kaydedildiler. Bu yapılandırılmış gözlemde, tacize uğramış ailelerin çocukları, kötü muamele görmemiş ailelerin çocuklarına göre daha az kendi başlarına oynadılar ve bu ailelerdeki anneler dikkat yönelimli davranışları bakımından

farklıydılar. Valentino ve arkadaşları, bu çalışmalarında kötü muamele içeren çevrenin çocukların sosyal-bilişsel gelişimleri üzerindeki etkilerini aydınlattılar ve nasıl müdahale edilebileceğini tartıştılar.

Gelişim psikologları sıklıkla yapılandırılmış gözlem kullanır. Jean Piaget'nin (1896-1980) bu yöntemi kullanış biçimi, belki de en dikkate değer olanıdır. Piaget'nin pek çok çalışmasında, çocuğa çözmesi için bir problem verilir ve daha sonra çocuğun anlayışının sınırlarını test etmek için problemin bir kaç farklı şekli daha verilir. Bu yapılandırılmış gözlem, çocuk zihni hakkında zengin bilgi sağlamıştır ve Piaget'nin zihinsel gelişimi açıkladığı "evre teorisi" için bir temel oluşturmuştur (Piaget, 1965).

Yakın zamanlarda, Evans ve Lee (2013) sade bir yapılandırılmış gözlem tekniği kullanarak çok küçük çocukların yalan söyleme davranışlarını incelediler. İki ve üç yaş çocuklarını bireysel olarak sessiz bir odada test ettiler. Araştırmacılar her çocuğa bir hediye paketi verdiler ama kendileri bu hediye için kurdele getirmeye gittikleri zaman gizlice bakmamasını istediler. Gizli kamera, araştırmacı gittikten sonra çoçuğun davranışlarını tespit etti. Toplam olarak çocukların %80'i hediyeye gizlice baktı. Ancak 2 yaşındakilerin çoğu baktığını itiraf etti, büyük çocukların (43-48 ay) %90'ı bakıp bakmadıkları hakkında yalan söylediler. Araştırmacılar, bilişsel gelişmenin artışıyla ortaya çıkan "yalan söylemenin, gelişimin ilk kilometre taşı olduğunu" öne sürdüler ve küçük çocuklar doğru söylemek konusunda ahlâken gelişmiş oldukları için değil, gerekli bilişsel işlevlerdeki yetersizlikleri nedeniyle, zorunlu olarak dürüst davranıyorlardı (s. 5).

Yapılandırılmış gözlem, pasif müdahalesiz doğal gözlem ile sistematik kontrolü olan ve bağımsız değişken müdahalesi yapılan laboratuvar deneylerinin tam ortasında bir yerdedir. Bu bağdaşma, araştırmacıların laboratuvar yerine daha doğal ortamlarda gözlem yapmalarına olanak sağlar. Yine de her şeyin bir bedeli vardır. Eğer gözlemci gözlem yaparken her seferinde aynı işlemleri izlemekte başarısız olursa, bu, aynı problemi inceleyen diğer araştırmacıların aynı sonuçları elde etmesini zorlaştırır. Kontrol edilemeyen ve belki de bilinmeyen değişkenler, gözlenen davranışın ortaya çıkmasında önemli bir rol oynayabilir. Bu sorunu ortadan kaldırmak için araştırmacılar işlem sürecinde tutarlı olmalı ve gözlem boyunca mümkün olduğunca benzer bir yapıyı gözlemeye çaba sarf etmelidir.

Temel Kavram

Alan Deneyleri Araştırmacı, davranış üzerindeki etkilerini belirlemek amacıyla doğal bir ortamda bir veya daha fazla bağımsız değişkeni değişimlediği zaman, **alan deneyi** yapmış olur. Alan deneyi, gözleme dayalı yöntemlerde müdahalenin uç bir örneğini sunar. Diğer gözleme dayalı yöntemlerle alan deneyi arasındaki temel fark, alan deneylerinde araştırmacıların bir bağımsız değişkeni değişimledikleri zaman kontrol için daha fazla çaba sarf etmeleridir. Alan deneyleri sıklıkla sosyal psikolojide kullanılır. Dikkat körlüğü üzerine çalışan araştırmacılar, çevredeki belirgin uyaran tiplerini değiştirmişlerdi, örneğin tek tekerlekli bisiklete binen palyaçonun etkisini kaykaya binen çocuk ile karşılaştırmışlardı. Bu şekilde, bir bağımsız değişkenin (uyaranın tipi) değişimlendiği çalışma, alan deneyi olarak isimlendirilmeye uygundur. Bir görsel ekleyerek

araştırmacıların bir süpermarkette yardım toplama kutusunun üzerinde yer alan resim çeşitlerini değişimledikleri bir alan deneyini ele alalım (Powell, Roberts ve Nettle, 2002). Kutuların üzerinde bulunan göz resmi, kontrol resmi (yıldız) ile kıyaslandığında, bağışları yaklaşık %50 artırmıştır. Araştırmacılar bu bulguyu, evrim boyunca insanların sosyal sonuçlarla bağlantılı ipuçlarına (göz) daha duyarlı hâle geldiğini söyleyerek, etolojik yaklaşıma göre açıklıyorlar (Bu bölümdeki Kutu 4.2.'ye bakınız). Bizim deneysel yöntem tartışmamız Bölüm 6'da devam edecek.

DOLAYLI (DİKKAT ÇEKMEYEN) GÖZLEME DAYALI YÖNTEMLER

- Dolaylı gözleme dayalı yöntemlerin en önemli üstünlüğü, tepkiselliğe neden olmamalarıdır.
- Dolaylı veya dikkat çekmeyen gözlemler, arşiv kayıtları veya fiziksel izlerin incelenmesiyle elde edilir.

Temel Kavram

Gözleme dayalı yöntemlerin gözlemcinin belirli bir ortamda davranışı doğrudan gözlemesi ve kayıt etmesi olduğunu söylemiştik. Bununla birlikte, davranış, insan davranışlarının diğer göstergeleri ve kayıtları kullanılarak da dolaylı olarak gözlenebilir. Bu yönteme çoğunlukla, **dikkat çekmeyen ölçümler** denilir çünkü araştırmacı duruma müdahalede bulunmaz ve birey gözlemin farkında değildir. Bu yöntemin üstünlüğü tepkiselliğin yokluğudur. Bir davranışsal ölçümde katılımcı, bir ölçme sürecini etkileyen gözlemcinin farkında olduğu zaman tepkisel davranır. Dikkat çekmeyen gözlemlerin dolaylı olması

İLERİ ALIŞTIRMA

Bu alıştırmada gözleme dayalı bir araştırmanın özet tanımlarını takip ederek cevaplayacağınız sorular soruyoruz.

Psikoloji yüksek lisans öğrencileri, üniversite kafeteryasında yemek yiyen heteroseksüel çiftler üzerinde gözleme dayalı bir araştırma yürütüyorlar. Çiftlerden hangisinin (kadın ya da erkek) elektronik cihaz (örneğin, telefon, tablet bilgisayar) kullanarak yemeği daha fazla kesintiye uğratacağını bulmak istiyorlar. Bu gözlemi yapmak için yardımcı araştırmacı olarak bir kaç lisans öğrencisi de onlara yardım ediyor.

Ekibin iki üyesi, öğlen ve akşam yemeklerinde birer saat, beş gün boyunca kafeteryada oturarak çiftleri gözlüyor. Gözlem, dört defa 15 dakikalık sürelerde yapılıyor. Kafeteryanın yanındaki merdiven yemek alanının gözlenmesine uygun ve gözlemciler basamaklara oturarak gözlemlerini kaydediyorlar. Gözlem süreci başladığında, her gözlemci 15 dakika boyunca veya onlar ayrılmak üzere kalkana kadar gözlemek için bir çift seçiyor. Gözlemciler, masada otururken kadın veya erkeğin hangisinin elektronik cihaz kullandığını işaretliyorlar.

Beş günlük gözlem sonunda, 15 dakikalık gözlemin tamamı veya bir kısmı için toplamda 80 çift gözlenmiştir. Bu 80 çiftin 60'ında, çiftlerin birisi veya ikisi birden elektronik cihaz kullanmıştır. Elektronik cihaz kullanılan 60 durumda, yalnızca çiftlerden birinin cihaz kullandığı 50 örnek (40 erkek, 10 kadın) vardı. Yüksek lisans öğrencileri, erkeklerin kadınlardan çok daha fazla elektronik cihaz kullanarak yemekleri böldükleri sonucuna vardılar.

1. Bu çalışmada yüksek lisans öğrencileri hangi gözlem yöntemini kullandı?
2. Örnek seçme işlemini nasıl betimlersiniz?
3. Sonuçların yüksek lisans öğrencilerinin vardığı sonuçları desteklediğini düşünüyor musunuz?
4. Bu çalışmada kullanılan gözlem yöntemi hakkında ne gibi çekinceleriniz var?

TABLO 4.1 DOLAYLI (DİKKAT ÇEKMEYEN) ÖLÇÜMLER

Fiziksel İzler	Arşiv Kayıtları
1. *İz kullanımı:* Bir maddenin kullanılması (veya kullanılmaması) sonucundaki fiziksel kanıt *Örnek:* geri dönüşüm kutusundaki konserveler, ders kitabında altı çizilmiş sayfalar, video oyunu kumandasındaki yıpranma ve aşınma	1. *Ardışık Kayıtlar:* aralıksız olarak oluşturulan kamusal ve özel belgeler *Örnek:* Spor takımlarının kayıtları, Facebook ve Twitter girişleri.
2. *Ürünler:* Eserler, yapılar veya diğer el emeği ürünü davranışlar *Örnek:* Tarih öncesi dönemlerde yapılmış taş üstü resimleri, MTV, *Harry Potter* oyuncakları	2. *Belirli Olayların Kayıtları:* Belirli olayları tanımlayan belgeler. *Örnek:* doğum belgeleri, evlilik belgeleri, okul dereceleri

Webb ve arkadaşları (1981) tarafından yapılan çalışma esas alınmıştır.

nedeniyle, araştırmacı gözlem yaparken insanların bir tepki vermesi veya davranışını değiştirmesi mümkün değildir.

Dikkat çekmeyen yöntemlerin aynı zamanda doğrudan gözleme dayanarak doğrulanabilen veya tartışılabilen önemli bilgiler sağlaması, bu yöntemi çok yöntemli (multimethod) araştırma yaklaşımlarının önemli bir aracı hâline getirir. Bu bölümde arşiv kayıtları ve fiziksel izlerin incelenmesini kapsayan bu dolaylı yöntemleri açıklayacağız.

Fiziksel İzler

- Fiziksel izlerin, "izlerin kullanımı" ve "ürünler" olmak üzere iki şekli vardır.
- İzlerin kullanımı, maddenin kullanılmasının (veya kullanılmamasının) fiziksel kanıtlarını ifade eder ve kullanım doğal hâlinde veya kontrollü koşulda ölçülebilir.
- Araştırmacılar, insanların kendi ürünlerini veya o ürünlerin üretildiği kültürü inceleyerek tutumlar, tercihler ve davranışlar hakkında hipotezler sınarlar.
- Fiziksel iz ölçümlerinin geçerliği, ortak bir noktada birleşen kanıtların aranmasıyla ve olası kaynak yanlılıkları gözetilerek incelenir.

Tanınmış bir yorumcuya, " bir el altı yemek kitabı"' olup olmadığı soruldu. Onun yoktu, fakat kocasının çok sevdiği bir tane olabileceğini söyledi, çünkü "sarımsak gibi kokardı ve yemek lekeleriyle doluydu" (The New York Times Book Review, 8 Nisan 2013, s. 8). Geçmiş davranışların fiziksel kanıtlarını incelemek olayların ve bireylerin tipik özellikleri hakkında önemli ipuçları sağlayabilir (Gosling, 2010). **Fiziksel izler** geçmiş davranışların kalıntıları, parçaları ve ürünleridir. İki tür fiziksel iz vardır; "kullanım izleri" ve "ürünler".

Temel Kavram

Kullanım izleri, isminden de anlaşılabileceği gibi, bir maddenin kullanılmasından (veya kullanılmamasından) kaynaklanan fiziksel kanıtlardır. Kül tablasındaki sigaradan kalanlar, geri dönüşüm kutusundaki teneke kutular ve kampüsün yürüyüş yollarındaki çer çöp kullanım izlerine birer örnektir. Saat ayarları, farklı kültürlerdeki insanların dakik davranmayı ne kadar önemsediklerini bize söyleyen bir kullanım izidir; ders kitaplarındaki izler öğrencilerin en çok hangi konuyu çalıştığı konusunda bilgi verir.

Ayrıca kullanım izlerini araştırmacıların belirli bazı maddelerin kullanımı yoluyla veri toplamaya müdahalede bulunup bulunmadıklarına göre

sınıflayabiliriz. *Doğal kullanım izleri*, araştırmacıların herhangi bir müdahalede bulunmadan gözledikleri bir olayın doğal oluşumudur. Tersine, *kontrollü kullanım izleri*, araştırmacının bazı müdahalelerinin bir sonucudur.

Araştırmacılar çevresel etmenlerin yiyecek alımını nasıl azaltabileceğini –dolayısıyla Amerikalılarda obezitenin nasıl düşürüleceğini– incelemek için bir dizi araştırma yaptılar ve bu çerçevede, araştırmanın bir parçası olarak dikkat çekmeyen ölçümler kullandılar. Kolej öğrencileri üzerinde yapılan bu çalışmada, öğrencilere film seyrederken yemeleri için patates cipsiyle doldurulmuş tüpler verildi (Geier, Wansink ve Rozin, 2012). Araştırmacılar bazı tüplere düzenli bir arayla (örneğin, her 7'nci cips) renkli cipsler koymuşlardı. Diğer tüplerde renkli cips yoktu. Filmin sonunda her öğrencinin tükettiği cips sayısı tespit edildi. Renkli cipslerin olduğu tüplerden yiyen öğrenciler yaklaşık olarak %50 daha az yediler.

Yenilen cips sayısının, bölünmemiş tüplerde öğrencinin yeme davranışının doğal kullanım izi olduğunu kabul edebiliriz. Tersine, renkli cipslerle bölünmüş tüplerde araştırmacı bir müdahalede bulunmuş olduğu için, yenilen cips sayısı kontrollü kullanım izidir. Araştırmacılar, bu sonuçlar için bir kaç muhtemel açıklama öne sürdüler. Bunlardan bir tanesi, bölümleme ipucunun (renkli cips) bir durma noktası işlevi görerek otomatik yeme davranışını kırmasıdır.

Ürünler, yaratılmış şeyler, yapılar veya diğer el emeği ürünü davranışlardır. Antropologlar sıklıkla antik kültürlerin günümüze kadar kalabilmiş ürünleriyle ilgilenirler. Boyamalar, kap-kacak, aletler ve diğer eserleri inceleyerek antropologlar, binlerce yıl önceki davranış örüntülerini tarif edebilirler. Çok sayıdaki modern zaman ürünü (örneğin, televizyon şovları, müzik, moda ve elektronik cihazları) kültürümüz ve davranışlarımız hakkında bir kavrayış sağlar. Örnek olarak, araç tamponlarına yapıştırılan çıkartmalar, halka açık duyguların ifadesi için bir çıkış yolu sunar ve aynı zamanda bireylere inançlarını ve özel gruplara aidiyetlerini ifade etme olanağı sağlar (Endersby ve Towle, 1996; Newhagen ve Ancell, 1995). Dövmeler ve hızmalar bazı kültürlerde benzer bir işlev görebilir (Swami, 2012; Bu bölümdeki Şekil 4.4' bakınız).

Ürünlerin incelenmesi, araştırmacılara davranışlar hakkında önemli hipotezleri sınama olanağı sunar. Örneğin psikologlar, "Fransız paradoksunu" sorgulamak üzere Amerika ve Fransa'daki besinlerle bağlantılı ürünleri incelediler (Rozin, Kabnick, Pete, Fischler ve Shields, 2003). "Fransız paradoksu" terimi, Fransızların Amerikalılara oranla daha fazla şişmanlatıcı yiyecek ve çok daha yağlı besin yemelerine rağmen, kalp hastalığından kaynaklanan ölüm ve obezite oranlarının Birleşik Devletleri'nkinden çok düşük olması demektir. Bu fark için bir kaç hipotez öne sürülmüştür; metabolizma farkı, kaygı düzeyi ve kırmızı şarap tüketimi. Rozin ve arkadaşları, Fransızların sadece az yediklerini iddia ettiler ve bu hipotezi sınamak üzere her iki ülkedeki besin ürünlerini, özellikle porsiyon büyüklüğünü incelediler. Onlar, Amerikan lokantalarında porsiyonların Fransız lokantalarıyla karşılaştırıldığında ortalama %25 daha büyük olduğunu ve Amerikan süper market raflarının ebadının genellikle daha geniş olduğunu buldular. Onların bu ürün gözlemi, kalp hastalıkları kaynaklı obezite ve ölüm oranındaki farkın nedeninin, Fransızların daha az yemeleri olabileceği hipotezini destekledi.

Fiziksel izler araştırmacılara, davranışın incelenmesi için değerli ve bazen yenilikçi bir yol sunar ve ölçümlerin işe yararlılığı, sadece araştırmacının becerisi ve kaynak bolluğu ile sınırlıdır. Bununla birlikte, fiziksel izlerin geçerliliği

ŞEKİL 4.4 Çoğu kültürde dövmeler ve hızmalar grup kimliği ve kendini ifade etmenin bir yolu olarak kullanılır.



ve bağımsız veri kaynaklarıyla doğrulanmalıdır. Geçerlik, ölçümlerin doğruluğunu ifade eder ve bütün ölçümlerde olduğu gibi fiziksel izlerde de insanların davranışları hakkında bizi doğru bir şekilde bilgilendirip bilgilendirmediğini sorgulamak zorundayız.

Kullanım izlerinin saklanma ve varlığını koruma şeklinde yanlılıklar olabilir. Örneğin, bir müze binasının sağ tarafındaki çok kullanılmış olan bir patika insanların ilgilerini çeken eserlerin yönünü mü gösterir, yoksa basitçe insanın doğal olarak sağa dönme eğilimini mi? Üniversitedeki geri dönüşüm kutularında bulunan teneke kutu sayısı öğrencilerin marka tercihleri midir, yoksa basitçe kampüs içindeki yiyecek otomatlarında mevcut olanlar mıdır? Amerika ve Fransa'daki süpermarket raflarının farklı büyüklüklerde olmasının nedeni ailelerdeki kişi sayısı mıdır, yoksa porsiyon büyüklüğü müdür? Araştırmacılar mümkün olduğu kadar fiziksel izlerin geçerliliği için destekleyici kanıtlar sağlamaya ihtiyaç duyarlar (Webb ve ark., 1981). Fiziksel iz gözleminde alternatif hipotezler göz önüne alınmalı ve sonuçlar diğer araştırmalarla karşılaştırılırken ölçümlerin benzer olduklarından emin olunmalıdır.

Arşiv Kayıtları

- Arşiv kayıtları, olayların tek veya ardışık kayıtlarını içeren ve bireylerin, grupların, kurumların ve devletin etkinliklerini betimleyen kamusal ve özel dokümanlardır.

- Arşiv verileri çok ölçümlü çalışmaların bir kısmını oluşturacak şekilde, doğal müdahalelerin etkilerini değerlendirmek, laboratuvar bulgularının dış geçerliğini belirlemek ve hipotezleri sınamak amacıyla kullanılır.
- Arşiv kayıtlarının potansiyel sorunları, seçici birikim, seçici kalıcılık ve sahte bağlantı olasılığıdır.

Bir anlığına sizin hakkınızda mevcut olan çok çeşitli kayıtları düşünün; doğum sertifikası, okul kayıt ve notları, kredi ve para kart harcamaları, ehliyet, çalışma ve vergi kayıtları, tıbbî kayıtlar, oy kullanma kayıtları, e-posta ve cep telefonu hesaplarınız ve eğer varsa Face-book, twitter gibi siteler... gündelik hayatınızı belirleyen sayısız kayıt. Şimdi bunu benzer kayıtları olan milyonlarca diğer insanlarınkiyle çarpın; oradaki bu sonsuz sayıda veri elinizin altındadır. Bütün bu verilere ulaşılması mümkün ülke, devlet, kurum, işyeri, medyayı ekleyin; arşiv kayıtları kullanarak insanların davranışlarını açıklayan psikologların kullandığı verilerin zenginliği kavramaya başladınız.

Temel Kavram

Arşiv kayıtları, bireylerin, grupların, kurumların ve devletlerin etkinliklerini tanımlayan kamusal ve özel kayıtlardır. Kayıtlar, sürekli saklanan ve güncellenen türde ise onlara ardışık kayıt denilir. Borsa veya spor takımlarının olduğu gibi, sizin akademik hayatınıza ilişkin kayıtlar (notlar, etkinlikler) birer ardışık kayıt örneğidir. Doğal ortamlardaki davranışlara bir pencere sağlayan Facebook girişleri de öyledir (ör. Wilson, Gosling ve Graham, 2012; www.facebookinthesocialscience.org). Bir araştırmacı, örneğin, facebook'ta milyonlarca "durum güncelleme"yi analiz ederek "ulusal mutluluk" üzerine çalıştı (Kramer, 2010). Facebook profilleri de kullanıcının sosyal kaygı gibi kişilik özelliklerini analiz etmek için kullanılabilir (Fernandez, Levinson ve Rodebaugh, 2012). Kişisel belgeler (örneğin, doğum belgesi, evlilik belgesi) gibi diğer kayıtlar, daha çok belirli durumu ya da olayı tanımlar ve onlara *olay kayıtları* denilir (Webb ve ark., 1981).

Bir davranış ölçümü olarak arşiv verileri, fiziksel izlerin bazı üstünlüklerini paylaşır. Her ikisi de anket, laboratuvar deneyi, doğrudan gözlem gibi hipotezleri sınamaya dayalı diğer yöntemleri tamamlayan ve dikkat çekmeyen ölçümlerdir. Çeşitli yaklaşımlardan gelen bulgular birbirleriyle uyumlu olduğu zaman, bulguların dış geçerliliği artar. Yani, bulguların farklı araştırma yöntemlerinin tamamına genellenebileceğini ve sınadığımız hipotezi desteklediğini söyleyebiliriz. Örneğin, Fransız paradoksunu açıklamak için porsiyon büyüklüğüyle ilgili fiziksel iz ölçümünü hatırlayın (Rozin ve ark., 2003). Bu araştırmacılar, aynı zamanda hipotezlerini sınamak için arşiv kayıtları üzerinde de çalışmışlardı. Onlar iki şehirde Filadelfiya ve Paris'te, lokanta tanıtım rehberlerini incelediler ve orada yazılı olan "açık büfelerin" sayısını kaydettiler. Mevcut bir arşiv kaydını kullanarak (lokanta tanıtım rehberleri), hipotezleriyle örtüşen kanıtlar elde ettiler. Filadelfiya'da 18 açık büfe yemek veren lokanta vardı fakat Paris'te hiç yoktu.

Araştırmacılar, *doğal müdahale*lerin etkilerini değerlendirmek için arşivleri inceleyebilir. Doğal müdahale, birey veya grupları ciddi şekilde etkileyen ve doğal olarak oluşmuş olaylardır. Bu olayları önceden tahmin etmek her zaman mümkün olmadığından, etkilerini değerlendirmek isteyen araştırmacılar arşiv verilerini içeren çeşitli davranış ölçümlerini kullanmak durumundadır. 11 Eylüldeki terör eylemi, 2008'deki dünya çapındaki ekonomik çöküş gibi ekonomik bir olay ve yeni yasa veya reformların yasalaşması, davranışlar

KUTU 4.3

GÖRÜNMEYEN EKONOMİ BİLİMİ

Öğretmenler kendileri ve öğrenciler daha iyi görünsün diye testlerde hilekârlık yapar mı?
Polis gerçekten düşük suç oranı ister mi?
Niçin ölüm cezası suçluları caydırmaz?
Çocuğunuz için hangisi daha tehlikelidir? Bir yüzme havuzunuzun olması mı yoksa bir silahınızın mı?
Niye doktorlar ellerini yıkamakta bu kadar kötüdür?
Terörist yakalamanın en iyi yolu nedir?
İnsanlar bencil ya da fedakâr olmaya mecbur mudur?
Eğer bu kadar etkisizse neden kemoterapi veriliyor?

Bunlar ve benzeri sorular, sıra dışı sosyal bilimci Steven D.Levitt'in çok satılan kitabında soruluyor (Freakonomics ve Super-Freakonomics, Levitt ve Dubner, 2005; 2009). Verdiği cevaplar, öğrencilerin test puanlarının analizi, spor kayıtları, suç istatistikleri, doğum ve ölüm istatistikleri ve daha bunlar gibi pek çok veriden geliyor. Bize yalnızca toplum arşivlerinde zekice maden arayarak bütün bu cevapların sırlarını vermiyor, aynı zamanda bize bu çok baskılı sınav çağında öğretmenlerin bazen hilekârlık yapabildiğini ve eğer hem silahınız ve hem de yüzme havuzunuz varsa, çocuğunuzun boğularak ölme olasılığının silahla ölme olasılığından 100 kez fazla olduğunu söylüyor.

üzerinde önemli etkileri olabilecek olaylara örnektir ve arşiv verileri kullanılarak incelenebilir. Aynı zamanda bireyler, ebeveyn ölümü ya da boşanması, kronik hastalıklar veya ilişkilerdeki zorluklar gibi hayatın içinde doğal olarak oluşan deneyimler yaşarlar. Bu tür olayların etkileri de arşiv verilerinden yararlanılarak araştırılabilir. Örneğin bir araştırmacı, çocukların ebeveyn boşanmasına verdiği tepkileri araştırmak için okula devamsızlık ve ders notlarını inceleyebilir.

Araştırmacılar arşiv kayıtlarını kullanarak pek çok pratik avantaj kazanabilirler. Arşiv verileri bereketlidir ve araştırmacı, geniş kapsamlı bir veri toplama aşamasından kurtulabilir; çünkü veriler araştırmacı için orada öylece durmaktadır! Aynı zamanda, arşiv bilgisi sıklıkla kamusal kayıtların bir parçası olduğu ve sıklıkla birey kimliği içermediği için etik endişeler daha az kaygı vericidir. Sayısız arşiv kaynağı internet aracılığı ile ulaşılabilir hale gelmiştir ve araştırmacılar bu şekilde davranışları incelemeyi daha zahmetsiz bulacaklardır (Kutu 4.3'e bakınız).

Bununla birlikte araştırmacılar, arşiv kayıtlarının sınırlılıklarının ve sorunlarının bilincinde olmalıdır. Seçici birikim ve seçici kalıcılık, bu veri toplama yönteminin iki önemli sorunudur [Webb ve arkadaşlarına (1981) bakınız]. Bu sorunlar arşivlerin oluşturulmasında yanlılık var olduğu için ortaya çıkar. *Temel Kavram*

Seçici birikim, bazı bilgilerin arşivde biriktirilmek üzere seçildiği ve bazılarının ise tercih edilmediği zamanlarda görülür. Örneğin, büyük bir arşivi düşünün; okul yıllıkları. Bütün etkinlikler, olaylar ve gruplar yıllığa konulmamıştır. Yıllıkta yer alanların daha öne çıkarılmasına kim karar vermiştir? Bazı olaylar, etkinlikler veya gruplar, diğerlerine oranla daha yüksek bir seçim şansı elde ettikleri zaman, bir yanlılık ortaya çıkar. Veya şunu düşünün: Politikacılar gibi sık sık gazetecilerle karşılaşanlar, bazı açıklamalarının "kayıt dışı" olduğunu söyleyerek medyayı kendilerine uygun şekilde yönlendirirler. Bu bir seçici birikim sorunu olarak görülebilir; çünkü sadece bazı bilgiler "kayıt için"dir. Burada bir tepkisellik sorunu da görebilirsiniz; şöyle ki, o açıklamayı yapanlar söylediklerinin "kayda alınacak" olmasına karar verdikleri zaman, açıklamalarının kaydediliyor olmasından etkilenirler, tepki verirler.

Facebook profillerini analiz eden araştırmacılar, bu bilgilerin, kullanıcının gerçek kişiliğini yansıtmak yerine bazen yanlış olabileceklerini veya "ideal ben"i temsil edebileceğini bilmelidir (örneğin, Back ve ark., 2010). Genel olarak pek çok kullanıcının kişisel bilgilerini belki biraz abartarak ve amaçlı şekilde düzenleyerek kendi benlik imgelerini yükseltmeye çabalayacakları beklenebilir (Walther ve ark., 2008).

Temel Kavram

Seçici kalıcılık, tamamlanmamış veya eksik kalmış kayıt olduğu zaman meydana gelir (araştırmacının farkında bile olmadığı bir şey). Araştırmacılar bazı kayıtların kalmış, diğerlerinin kalmamış olabileceğini göz önüne almak durumundadır. Belgelere belirli bir birey veya grup tarafından özellikle hasar verilmiş veya tamamen yok edilmiş olabilir; örneğin, bir başkanlık yönetiminden diğerine geçiş sürecinde böyle olabilir. Aynı şekilde, aile fotoğraf albümlerinden "gizemli" bir şekilde fotoğraf kaybolabilir; örneğin, boşanmış olanların evlilik dönemlerine ait veya şişmanlık zamanlarına ait olan fotoğraflar bu akıbete uğrayabilir. Dergi ve gazete köşe yazarları ve editörleri aldıkları bir mektubun yalnızca bir kısmını basarlar; böylece, sadece o kısmı "kalıcı" olur.

Arşiv verilerinin analizinde bir diğer sorun, *sahte ilişki*lerin bulunması olasılığıdır. Bir sahte ilişki, bir ya da iki değişkenin bağlantılı olduğunu gösteren hatalı kanıtlar varsa ortaya çıkar (Bölüm 5'e bakınız). Hatalı kanıtlar, uygun olmayan veya yetersiz istatistik analizler ya da sıklıkla bir değişken tesadüfen diğerleriyle ilişkili göründüğü zaman ortaya çıkar. Çoğunlukla belirlenememiş olan üçüncü bir değişken ilişkiyi açıkladığı zaman, iki değişken arasında bir bağlantı veya ilişki ortaya çıkabilir. Örneğin, arşiv kayıtları dondurma satışlarıyla suç oranlarının bağlantılı olduğunu gösteriyor (yani, dondurma satışları arttıkça, suç oranları da artıyor) olsun. Dondurma yemenin insanların suç işlemesine neden olduğuna karar vermeden önce, her iki değişkeni –dondurma satışları ve suç oranları– birlikte ele alarak muhtemelen mevsim ısısı gibi üçüncü bir değişkenden etkilendikleri düşünülmelidir. Dondurma satışları ve suç oranları arasındaki sahte ilişki, üçüncü bir değişken olan ısı ile açıklanabilir.

Sahte ilişkilerde olduğu gibi, seçici birikim ve seçici kalıcılık nedeniyle meydana gelen yanlılık olasılığı da araştırmacıların sadece arşiv çıktılarına dayanarak sonuca varmalarının yanlış olabileceği anlamına gelir. Arşiv kayıtları, olguların çoklu ölçüm yaklaşımıyla araştırıldığı durumlarda, ek kanıtlar sağlamak için kullanılırsa, daha yararlı olur.

DAVRANIŞIN KAYDEDİLMESİ

- Gözleme dayalı araştırmaların amacı, araştırmacının davranışın bütününü mü yoksa belirli bir kısmını mı araştıracağını belirler.
- Sonunda araştırma sonuçlarının nasıl özetleneceği, analiz edileceği ve rapor edileceği, başlangıçta davranışların nasıl gözlendiğine ve kaydedildiğine bağlıdır.

Doğrudan ve dolaylı gözlemlere ek olarak, gözleme dayalı yöntemler davranışın nasıl kaydedildiğine bağlı olarak da farklılaşır. Bazen araştırmacılar davranışın ortaya çıktığı ortamın ve davranışın *kapsamlı* bir açıklamasını yapmak isterler. Daha sık olarak ise yalnızca belirli bir olay ve davranışa odaklanırlar. Yalnızca *seçilmiş* davranışların mı yoksa bir ortamdaki bütün davranışların mı gözleneceği, araştırmacıların amacına bağlıdır. Kaydedilmek için hangi davranışın seçildiği,

sonunda hangi sonuçların ölçüleceğini, özetleneceğini, analiz edileceğini ve rapor edileceğini belirler.

Kapsamlı Davranış Kayıtları

- Davranışın yazılı olarak açıklandığı anlatı kayıtları ile işitsel ve görsel kayıtlar, kapsamlı kayıtlardır.
- Araştırmacılar davranış hakkındaki hipotezlerini sınamak için anlatı kayıtlarından elde edilen veriyi düzenler ve sınıflar.
- Anlatı kayıtları, gözlenen davranıştan hemen sonra veya gözlem esnasında yapılmalıdır ve gözlemciler, davranışları belirlenen ölçütlere göre kaydetmek için titiz bir şekilde eğitimden geçirilmelidir.

Temel Kavram

Araştırmacılar, davranışın kapsamlı bir kaydının peşinde oldukları zaman, sıklıkla anlatı kayıtları kullanırlar. **Anlatı kayıtları**, davranışın orijinal hâline az veya çok sadık bir kopyasını sağlar. Bir anlatı kaydı yaratmak için gözlemci, davranış betimlemelerini yazabilir veya işitsel ya da görsel olarak kaydedebilir. Örneğin, bir araştırmada tacize uğramış veya uğramamış ailelerde anne-çocuk etkileşimini kaydetmek için video kullanıldı (Valentino ve ark., 2006). İnternet sohbet odalarında katılımcı ergenlerin kimlik ve cinsiyet konusunu incelemek için araştırmacılar, iki aylık bir dönem boyunca kaydedilen oturumlardan seçkisiz olarak seçilmiş 30 dakikalık bir oturumdaki internet konuşmalarını incelediler (Subrahmanyam ve ark., 2004). Ergen sohbetlerinin yazıya dökümü 19 sayfa (815 satır) tuttu. İnternette davranışların kaydını otomatik olarak yapmanın yolları ve elde edilen verinin niteliksel analizine yardım edecek bilgisayar programları vardır [Utz'a (2010) bakınız].

Bir anlatı kaydı oluşturulduktan sonra araştırmacılar, incelenen davranış hakkındaki hipotezlerini sınamak için bu kaydı sınıflar ve düzenler. Anlatı kayıtları, diğer davranış ölçümleri ve kayıt şekillerinden farklıdır, çünkü davranışın sınıflandırılması gözlemden *sonra* yapılır. Bu nedenle araştırmacılar, anlatı kaydının araştırma hipotezini değerlendirmek için gerekli bilgiyi yakaladığından emin olmalıdır.

Anlatı kayıtları, ne her zaman gözlemcinin izlenim ve çıkarımlarını engeller ne de her zaman davranışın kapsamlı bir betimlemesi anlamına gelir. Örneğin, *alan notları* yalnızca gözlemcinin ilgilendiği katılımcı, olay, ortam ve davranışların ardışık bir dökümünü içerir ve mevcut her şeyin tam ve hatasız bir kaydından oluşmaz. Alan notları gazeteciler, sosyal hizmet uzmanları, antropologlar, psikologlar vb. tarafından kullanılır ve muhtemelen diğer anlatı kayıtlarından daha sık kullanılmaktadır. Olaylar ve davranışlar muhtemelen gözlemcinin uzmanlık bilgisi çerçevesinde yorumlanmış olmaktadır ve oldukça kişisel olmaya eğilimlidir (Brandt, 1972). Örneğin, bir klinik psikolog, birisine konulmuş bir teşhisi ve bazı klinik meseleleri bilen birisi olarak, o insanın belirli bazı davranışlarını kaydetmiş olabilir. Alan notlarının bilimsel bir kayıt olarak faydası, içeriğinin doğruluğu ve hassasiyetine bağlıdır. O hâlde, özellikle gözlemcilerin eğitilmesine ve kayıt edilmiş gözlemlerin bağımsız gözlemciler tarafından ne ölçüde doğrulandığına ve araştırmanın diğer olanaklarına bağlıdır.

Pratik ve yöntemsel etmenler, anlatı kayıtlarının ne şekilde yapılacağını belirler. *Genel bir kural olarak kayıtlar, davranış gözlenirken veya davranış gözlendikten*

mümkün olduğu kadar kısa bir süre sonra yapılmalıdır. Zamanın akışı ayrıntıları bulanıklaştırır ve eylemin orijinal sırasını tekrar oluşturmak gerekir. Bir anlatı kaydının neleri kapsayacağına, gözlemcinin hangi ölçüde çıkarım yapacağına, ve kaydın ne zaman tamamlanacağına, çalışma başlamadan önce karar verilmiş olmalıdır. Önceden anlatı kaydının içeriğine karar verildiyse, gözlemci belirlenmiş ölçütlere göre davranışı kaydetmek için eğitimden geçirilmelidir. Gözlem pratiği yapılabilir ve "gerçek" veri toplanmadan önce, bu kayıtlar birden fazla araştırmacı tarafından eleştirilebilir.

Seçilmiş Davranış Kayıtları

- Araştırmacılar belirli bir olay veya davranışı araştırdıkları zaman, çoğu kez davranışın sıklık ya da süre gibi niceliksel bir ölçümünü elde ederler.
- Davranışın niceliksel ölçümü kategorik, sıralı, aralıklı ve oranlı olmak üzere dört ölçüm düzeyinden birisinde yapılır.
- Çoğu kez psikolojik boyutları ölçmek için kullanılan dereceleme ölçekleri, genellikle sıralama düzeyinde oldukları hâlde, aralıklı ölçek olarak kabul edilirler.
- Elektronik kayıt cihazları doğal ortamlarda davranışı kaydetmek için kullanılabilir ve çağrı cihazları bazen, davranışını (örneğin, bir ankete) bildirmesi için haber verir.

Çoğu kez araştırmacılar ortam ve bireylerin bazı yönleri veya yalnızca belirli bir davranış ile ilgilenirler. Bekledikleri davranış konusunda belirli bir hipotezleri ve inceledikleri davranışın net bir betimlemesi olabilir. Gözleme dayalı araştırmaların bu şeklinde araştırmacılar, tipik olarak, gözlem yaparken belirli bir davranışın oluşumunu ölçerler. Örneğin, dikkat körlüğü çalışmalarında Hyman ve meslektaşları (2009) insanların palyaçoyu fark edip etmedikleri davranışları seçtiler ve palyaçoyu fark eden ve fark etmeyen insanların sayısını belirlediler. Yani onlar için "dikkat körlüğü"nün işevuruk tanımı, palyaçonun fark edilip edilmediğiydi (İşevuruk tanım tartışması için Bölüm 2'ye bakınız).

Doğal gözlem kullanarak insanların fiziksel engelli bireylere tepkilerini gözlediğinizi düşünün. İlk olarak, "fiziksel olarak engelli kişi"yi ve böyle bir kişiye verilecek bir "tepki"yi işevuruk olarak tanımlamanız gerekir. Yardım davranışlarıyla mı ilgileniyorsunuz, yaklaşma-kaçınma davranışlarıyla mı, göz teması kurma, sohbetin uzunluğu veya bir başka fiziksel tepki ile mi ilgileniyorsunuz? Bu durumda, bu davranışları nasıl ölçeceğinize karar vermeniz gerekir. Varsayalım ki, fiziksel engeli olan ve olmayan bireyler arasındaki göz temasını gözleyerek, insanların tepkilerini ölçmeyi tercih ettiniz. Göz teması kurmayı tam olarak nasıl ölçeceksiniz? Basitçe, bireylerin göz teması kurmalarını ya da kurmamalarını mı ölçeceksiniz, yoksa herhangi bir göz temasının süresini mi ölçeceksiniz? Kararınız, çalışmanızın hipotezine veya amaçlarına bağlıdır ve aynı ya da benzer davranışsal ölçümleri kullanan önceki çalışmalardan edindiğiniz bilgilerden etkilenecektir.

TABLO 4.2 ÖLÇÜM DÜZEYLERİNİN ÖZELLİKLERİ

Ölçek Tipi	İşlemler	Hedef
Kategorik	Eşit/eşit değil	Uyaranı somut kategorilere ayırmak
Sıralı	Daha büyük/daha küçük	Uyaranı tek bir boyutta sıraya dizmek
Eşit aralıklı	Toplama, çıkarma, çarpma, bölme	Belirli bir boyutta uyaranlar arasındaki mesafeyi belirlemek
Oranlı	Toplama, çıkarma, çarpma, bölme, oranları bulma	Verilen bir boyutta uyaranlar arasındaki mesafeyi belirlemek ve ölçek değerlerinin oranlarını göstermek

Temel Kavram

Ölçüm Düzeyleri Araştırmacılar belirli bir davranışın miktarını belirlemeye ve ölçmeye karar verdikleri zaman, hangi ölçüm düzeyini kullanacaklarına da karar vermelidir. Ölçümlerin hem fiziksel ve hem de psikolojik ölçümlere uygulanabilen dört seviyesi veya **ölçüm düzeyi** vardır: Kategorik, sıralı, eşit aralıklı ve oranlı. Her bir ölçüm düzeyinin özellikleri Tablo 4.2'de belirtilmiştir ve ayrıntılı tanımlamaları Kutu 4.4'te verilmiştir. Araştırma sonuçlarını analiz etmek için istatistik işlemleri seçerken bu dört ölçüm düzeyini aklınızda tutmanız gerekecektir. Verilerin nasıl analiz edileceği, kullanılan ölçüm düzeyine bağlıdır. Bu bölümde, gözleme dayalı araştırmalarda kullanılan ölçüm düzeylerini açıklayacağız.

Bir *kontrol listesi* çoğu zaman kategorik ölçümleri kaydetmek için kullanılır. Gözlemci basitçe, belirli bir davranışın ortaya çıkıp çıkmadığını "işaretler". Kontrol listelerinde sıklıkla, katılımcının ırk, cinsiyet ve yaşı gibi özelliklerini olduğu kadar, günün zamanı, yeri gibi ortamın özelliklerini ve bir başka insanın olup olmadığına ilişkin gözlemleri kaydetmek için bir boş alan da bulunur. Araştırmacılar tipik olarak, bu katılımcı ve bağlam değişkenlerinin bir fonksiyonu olarak gözlenen davranışla ilgilenirler. Örneğin, Hyman ve ark. (2009) dikkat körlüğü çalışmalarında yayaları yalnız veya çift olarak yürüyenler ile cep telefonu veya müzik aleti kullananlar olarak 4 kategoriye ayırdı (3 veya daha fazla kişinin birlikte yürüdüğü grupları ise hariç tuttuklarına dikkat ediniz).

Kutu 4.4'te belirtildiği gibi, ölçmenin ikinci düzeyi, gözlemlerin sıralanması ve sıraya dizilmesini kapsayan sıralama ölçekleridir. Örneğin bir araştırmacı, katılımcılara çeşitli uyaranları (nesne, resim, kişi) tercih veya çekicilik gibi bazı boyutlara göre nasıl sıralayacaklarını sorabilir. Katılımcı çocukların üç bebeği en çok sevdiklerinden en az sevdiklerine doğru sıralamalarını isteyen bir çocuk psikoloğu, sıralama düzeyinde bir ölçüm yapmış olur.

Gözleme dayalı araştırmada davranışı ölçmek için gözlemci, bazen koşulların ve özelliklerin miktarı veya düzeyi hakkındaki öznel kararına dayanarak, davranışın veya olayın derecesini belirler. Tablo 4.3'te eğitilmiş gözlemcilerin kullandıkları, annelerin çocuklarına yönelik sıcaklık ve duygusal yakınlık düzeylerini ayırt eden bir ölçek görüyorsunuz (Dickie, 1987). Yedi dereceli bu ölçeğin şu özelliklerine dikkat ediniz. İlk olarak birinci düzey, sıcaklık ve duygusal yakınlığın yokluğunu veya çok az olmasını temsil etmektedir; artan sayılar, bu özelliklerin miktarının da arttığını gösterir. İkinci olarak, gözlemcinin düzey belirlemesine yardımcı olmak üzere, tek sayılı ölçek değerleri (1, 3, 5, 7)

KUTU 4.4

ÖLÇÜM DÜZEYLERİ

En düşük ölçüm düzeyi, olayların bir kaç ayrı kategori altında sınıflanmasıdır ve kategorik ölçüm olarak adlandırılır. Örneğin, insanların göz renklerini "kahverengi gözlüler" ve "mavi gözlüler" şeklinde sınıflayarak ölçebiliriz. İnsanların fiziksel engelli bireylere tepkilerini incelerken araştırmacılar, katılımcıları göz teması kuranlar ve kurmayanlar olarak kategorize ederek sınıflayıcı ölçüm kullanabilirler.

Kategorik ölçümlerde verilerin özetlenmesi ve analizi için kullanılabilecek işlemler kısıtlıdır; kullanabileceğimiz tek aritmetik işlem, "eşit" ile "eşit değil" arasındaki ilişkiye dayanır. Kategorik verileri özetlemenin yaygın bir yolu, frekansları oran olarak göstermek veya kategorilerin yüzde değerlerini vermektir.

Ölçümlerin ikinci düzeyi, sıralı ölçümdür. Sıralı ölçüm, ölçülmüş durumların sıralanması veya derecelendirilmesidir. Sıralı ölçümler "daha büyüktür" veya "daha küçüktür" şeklindeki aritmetik ilişkileri de işe katar. Yarış sonuçları aşina olduğumuz bir sıralı ölçümdür. Bir koşucunun gümüş madalya kazandığını duyduğumuzda, onun ikinci olduğunu biliriz ancak onun kaç saniye sonra bitiş çizgisine ulaştığını veya altın madalyalı koşucunun kaç metre arkasından yarışı tamamladığını bilemeyiz.

Ölçümlerin üçüncü düzeyi, eşit aralıklı ölçümdür. Eşit aralıklı ölçüm, belirli bir boyutta iki durum arasındaki mesafenin belirlenmesine imkân verir. Sıralı ölçümlerde birinci ve üçüncü sıra arasındaki mesafe, üçüncü ve beşinci sıra arasındaki mesafeye eşit olmak zorunda değildir.Örneğin yarışı birinci ve üçüncü bitiren koşucular arasındaki süre farkı, üçüncüyle beşinci arasındaki süre farkına eşit olmayabilir. Bununla birlikte eşit aralıklı ölçümlerde, sayısal değerler arasındaki fark eşittir. Örneğin bir yetenek testinde 50 ve 70 doğru cevap arasındaki fark, 70 ile 90 doğru cevap arasındaki farka eşittir. Fakat eşit aralıklı ölçümlerde bir mutlak sıfır noktası yoktur. Örneğin eğer bir birey sözel yetenek testinden sıfır puan almışsa bu, o bireyin sözel yeteneğinin hiç olmadığı anlamına gelmez (muhtemelen bu bireyin, en sonunda bu testi geçebilecek kadar bir sözel kapasitesi vardır). Eşit aralıklı ölçümlerde toplama, çıkarma, çarpma ve bölme gibi standart aritmetik işlemler yapılabilir. Psikologlar, mümkün olduğunca psikolojik boyutları en azından eşit aralıklı düzeyde ölçmeye çaba sarfederler.

Ölçümlerin dördüncü seviyesi oranlı ölçümlerdir. Oranlı ölçümler eşit aralıklı ölçümlerin bütün özelliklerine sahiptir. Ama bu ölçekte, aynı zamanda bir mutlak sıfır noktası vardır. Aritmetik işlemler açısından bir sıfır noktası, oranlı ölçekleri anlamlı kılar. Süre, ağırlık, mesafe ölçen fiziksel ölçekler genellikle oranlı ölçek olarak kabul edilir. Örneğin 200 kg ağırlığındaki birisi, 100 kg. ağırlığındaki birisinin iki katıdır.

için açık ve net sözel tanımlamalar yapılmıştır. Çift sayılı değerler (2, 4, 6), gözlemci tarafından tek sayılı tanımlamaların arasına düşen davranışları derecelemek için kullanılmaktadır. Yedinci düzeyden daha ilerisi yoktur. Araştırmacılar farklı bitiş noktası olan ölçekler kullanırlar fakat 10 dereceli ölçekten fazlası genellikle kullanılmaz.

İlk bakışta, Tablo 4.3'te gösterilene benzer dereceleme ölçekleri, eşit aralıklı ölçek gibi görünür –gerçek bir sıfır noktası yoktur ve sayılar arasındaki aralıklar eşittir. Bununla beraber, yakından incelendiğinde, görülmektedir ki gözlemcilerin kullandığı pek çok dereceleme ölçeği, insanları veya olayları, psikolojik boyutlar üzerinde aslında yalnızca sıralama bilgisi sağlayacak şekilde değerlendirir. Bir dereceleme ölçeği olarak aralıklı düzeyde olmak örneğin, ikinci derece ile üçüncü derece

TABLO 4.3 EBEVEYN-ÇOCUK ARASINDAKİ SICAKLIĞIN VE DUYGUSAL YAKINLIĞIN ÖLÇÜMÜNDE KULLANILAN DERECELEME ÖLÇEĞİ.*

Ölçek Değeri	Açıklama
1	Sıcaklık, duygusal yakınlık ve hoşnutluğun yokluğu. Aşırı öfke, soğukluk, mesafe ve çoğunlukla çocuktan uzak durma. Saldırı düzeyinde ilişki.
2	
3	Zaman zaman sıcaklık ve etkileşimden hoşnutluk var. Ebeveynin çocuğuyla övündüğüne dair çok az belirti var veya bu övünç alışılmadık, tuhaf bir şekilde gösteriliyor. Ebeveynin ilişki şekli yapmacık ve tasarlanmış; samimi değil.
4	
5	Ortalama düzeyde hoşnutluk ve çocukla etkileşimde sıcaklık var. Ebevyn bazı konularda memnun, ama bazılarında değil.
6	
7	Sıcaklık ve hoşnutluk, çocukla etkileşimin temel özelliği. Çocuktan övünç ve hoşnutluk duyulduğuna dair yeterli belirti var. Hoşnutluk tepkileri çocuğun davranışlarına uygun.

*Dr. Jane Dickie'den alınmıştır.

arasındaki mesafenin 4 ile 5, 5 ile 6, 6 ile 7 arasındaki mesafeye eşit olmasıdır. İnsanı gözleyenler sıcaklık, hoşnutluk, saldırganlık veya kaygı gibi özellikler için öznel kararlar verebilirler ve dereceler arasındaki mesafeye dakik aralıklar verilmesi hemen hemen olanaksızdır. *Bununla beraber, pek çok gözlemci dereceleme ölçeği kullandıkları zaman aralıklı seviyede bir ölçüm yaptığını farz eder.* Herhangi bir davranış için uygulanan ölçüm düzeyine karar vermek her zaman çok kolay değildir. Eğer emin değilseniz, konunun uzmanlarının önerilerini alın; böylece verilerinizin betimlenmesi ve istatistiksel analizi konusunda uygun kararlar verebilirsiniz.

Kontrol listeleri, bir bireyin ya da bir grubun belirli bir zaman içinde belirli bir davranışının *sıklığını* ölçmek için de kullanılabilir. Önceden belirlenmiş her gözlem süresi içinde, belirli bir davranışın meydana gelip gelmediği işaretlenir. Bütün gözlemler tamamlandıktan sonra, araştırmacılar belirli bir davranışın kaç kez meydana geldiğine bakarlar. Bu durumda, verilen tepkilerin sıklığı, oranlı ölçme seviyesi olarak kabul edilebilir. Çünkü eğer "birim" davranış (örneğin, bir çocuğun sandalyesinden kalkması) sayılmış ise, bu durumda "sıfır", belirli bir davranışın hiç olmadığını gösterir. Oranlı ölçek değerleri anlamlı da olabilir. Örneğin, Sandalyesinden 20 kez kalkan bir çocuk, bu hedef davranışı, sandalyesinden 10 kez kalkan bir çocuktan iki kat daha fazla göstermiştir.

Elektronik Kayıt Alma ve İzleme Davranış elektronik kayıt alma ve izleme cihazları kullanılarak da ölçülebilir. Örneğin, kolej öğrencilerinde kan basıncı ile bilişsel baş etme stratejileri arasındaki ilişkileri inceleyen bir araştırmada katılımcılar, birisinde sınav olan iki "tipik" okul gününde, üzerlerine bir kan basıncı göstergesi taktılar (Dolan, Sherwood ve Light, 1992). Katılımcılar aynı zamanda kendi gündelik etkinlikleri ve baş etme stratejileri hakkında

birer soru listesi de cevapladılar. Araştırmacılar baş etme tarzlarının bir fonksiyonu olarak, günün farklı zamanlarında okudukları kan basıncı değerlerini karşılaştırdılar. "Yüksek kendi-odaklı baş etme" stratejisi (örneğin, "stresli durumlarda kendini tut/veya suçla" gibi s. 233) kullanan öğrencilerin kan basıncı, kendi odaklı baş etme stratejisi kullanmayanlara kıyasla, sınav esnasında ve sonrasında daha yüksek oldu.

Bir diğer elektronik yöntem, katılımcıların her gün güvenli bir internet sitesine (e-posta hatırlatmasıyla) girerek günlük olayları yazdıkları "internet günlüğü"dür. Kolej öğrencilerinin duygu durumları ile baş etme tarzlarının araştırıldığı bir çalışmada öğrenciler, her gün kendilerinde endişe uyandıran olayları ve onunla nasıl baş ettiklerini bildirdiler (Park, Armeli ve Tennen, 2004). Bu araştırmanın sonuçları gösterdi ki olumlu duygu durumunda, problem odaklı baş etme stratejisi, kaçınma odaklı baş etme stratejisinden daha fazla benimseniyor; bu durum, özellikle strese neden olan olayların kontrol edilebileceği düşünüldüğü zaman böyledir. Başka araştırmacılar, katılımcılara taşınabilir bir bilgisayar verdiler ve kendilerine hatırlatma geldiği zaman "elektronik günlüğe" yazmalarını istediler (örneğin, McCarthy, Piasecki, Fiore ve Baker, 2006; Shiffman ve Paty, 2006). Elektronik kayıt alma ve izleme yöntemleri, katılımcının duygu durumunu ve etkinliklerini bildirmesine dayanır ve davranışın doğrudan bir gözlemi değildir. Böyle veri toplarken olası yanlılıkları gözden kaçırmamak gerekir (örneğin, bazı etkinliklerin atlanması, hatalı ifade edilme olasılığı; bu gibi olası bazı yanlılıklara ilişkin tartışmalar için Larson, 1989'a bakınız).

GÖZLEME DAYALI VERİLERİN ANALİZİ

- Araştırmacılar gözleme dayalı verileri özetlemek için niceliksel veya niteliksel veri analizini seçerler.

Davranışlara ait gözlem verileri kaydedildikten sonra araştırmacılar, insanların davranışlarını özetlemek için bu gözlem verilerini analiz ederler ve bu gözlemlerinin güvenirliğini belirlerler. Araştırmacının seçeceği veri analizinin türü, araştırmanın amacına ve verinin nasıl toplandığına bağlıdır. Örneğin, araştırmacılar bir dereceleme ölçeği kullanarak seçilmiş bazı davranışları kayıt etmişlerse, niceliksel veri analizi (istatistiksel özetleme ve analiz) tercih edilir. Kapsamlı anlatı ve arşiv kayıtları varsa, hem niteliksel hem de niceliksel analizleri seçebilirler. Biz önce niteliksel analizleri betimleyeceğiz.

Niteliksel Veri Analizi

- Arşiv kayıtları ve anlatıların analizinde en önemi aşama verilerin indirgenmesidir.
- Araştırmacılar belirli ölçütlere (örneğin kategorilere) göre davranışı kodlar.
- Arşiv kayıtları ve anlatıların incelenmesinde içerik analizi kullanılır ve üç aşamadan oluşur: İlgili kaynakların belirlenmesi, kaynaktan örneklemler seçilmesi ve analiz birimlerinin kodlanması.

Gözleme dayalı araştırmalarda, çoğunlukla sayfalarca veya fazlasıyla uzun görsel ve işitsel malzemeden oluşan ve zengin bilgi sağlayabilen kapsamlı davranış kayıtları elde edilir. Bu bilgi bloglar, sosyal paylaşım siteleri ve diğer

internet kayıtlarından (örneğin, You Tube) sağlanan arşiv kayıtlarını da kapsayabilir. Esasen, iletişim türlerinin her şekli (televizyon ve radyo programları, filmler, konuşmalar, görüşmeler veya "twit"ler) analize açıktır.

Niteliksel veri analizi bir kaç farklı şekilde olabilir. Niteliksel araştırmalar tipik olarak temaları (konuları) ve davranış kategorilerini belirleyerek ve böylece özet veya sözel bir yorum elde etmek suretiyle anlatı kayıtlarındaki anlamı ararlar. Nitel araştırmalarda araştırmacılar, çoğu zaman araştırmalarına bir hipotezle başlamaz; bunun yerine, analizlerinden hipotez üretirler. Eğer sayılar kullanılmışsa, bunlar basit istatistiksel tablolarla verilir (niteliksel veri analizinin diğer özellikleri ve bunlar hakkındaki tartışmalar için örneğin, Silverman, 2011'e bakınız).

Temel Kavram

Bir anlatı kaydındaki bilginin miktarı yıldırıcı olabilir ve genellikle bazı **veri indirgeme** biçimlerinin uygulanmasını gerektirir. Böylece davranış verilerinin özüne ulaşılabilir ve özetlenebilir. Niteliksel veri analizinde araştırmacı, bilgileri sözel olarak özetlediğinde, temaları belirlediğinde, bilgiyi grupladığında ve bu anlatı kayıtları hakkındaki kendi gözlemlerini dâhil ettiğinde veri indirgeme yapılmış olur. Araştırmacılar bir kavrayışa ulaşmak için hem tümevarım ve hem de tümden gelime dayalı muhakeme kullanırlar. Bunu yaparken kendi kişisel geçmişlerinin kayıtların yorumunu etkileyeceğinin de daima farkında olmalıdırlar (Creswell, 2013).

İster niteliksel, isterse niceliksel analiz kullansınlar, araştırmacılar anlatı kayıtlarını analiz edecekleri zaman üç önemli sırayı izlerler: Uygun kaynakların tespit edilmesi, kaynaktan uygun örneklerin seçilmesi ve analiz biriminin kodlanması. Uygun kaynakların seçilmesi, araştırmacıların araştırma sorusunu cevaplamalarına olanak sağlayan bir hususdur. Bu süreç nispeten doğrudan olabilir. Örneğin, İkinci Bölüm'de sözünü ettiğimiz Kidd ve Kral'ın (2002) yaptığı niteliksel araştırma, sokak çocuklarının deneyimlerini anlamaya çalışıyordu. Bu örnekte araştırma sorusu, uygun kaynağı açıkça belirliyordu: Sokakta yaşayan genç insanlar. Bazı durumlarda kaynağın seçilmesi bir miktar beceri gerektirir; tıpkı araştırmacıların, ergenlerin kimlik ve cinsellik ile ilgili kaygılarını ortaya koymak amacıyla internet sohbet odalarındaki karşılıklı konuşmaları gözlerken yaptıkları seçimde olduğu gibi (Subrahmanyam ve ark., 2004).

Niteliksel analizlerde ikinci aşama, kaynaktan uygun bir şekilde örneklerin seçilmesini içerir. Pek çok arşiv kaydı, kaynaktaki bütün bilgilerin analiz edilmesini imkânsız hâle getirecek kadar yüklüdür. Böylece araştırmacı, yapacağı seçimin, örneklemin temsil ediciliğini etkileyeceğinin farkında olarak verinin bir kısmını seçmek zorundadır. Niteliksel araştırmacılar bir kaynağı sıklıkla ilgilerine veya uygunluğuna göre seçerler (Silverman, 2011). Örneğin, online sohbet odalarındaki ergen söylemlerinin analizini yapmak için araştırmacılar, iki aydan uzun bir süre izleyip kaydettikleri bir dizi oturumdan örnek seçtiler (Subrahmanyam ve ark., 2004).

Temel Kavram

Kaynağın analiz edilmesinde üçüncü aşama olan **kodlama,** çalışmanın amacıyla bağlantılı olacak şekilde, davranış birimlerinin belirlenmesidir (kategoriler ve temaları içerir). **İçerik analizi,** araştırmacıların kaydın belirli niteliklerine dayanarak çıkarımlar yapmasına izin veren herhangi bir nesnel kodlama tekniği olarak tanımlanır (Holsti, 1969). Kodlama, internet sohbet odaları katılımcılarının kullandıkları takma isimlerin yazılı dökümlerini belirlemede veya belirli davranışları adlandırmada olduğu gibi, nispeten basit olabilir. Ama, internet blogları gibi yazılı anlatıların niteliksel veri analizi, hem materyaldeki bütün temaları

Temel Kavram

hem de yazarın ifade ettiği duyguları ve ihtiyaçları kodlamayı amaçlar [örneğin, Mazur, 2010; niteliksel veriler analiz edilirken kullanılan kodlama stratejilerine genel bir bakış için, Saldana'ya (2013) bakınız].

Kaynak seçiminde olduğu gibi, araştırmanın amacı, kodlanacak temaları ve betimsel kategorileri belirlemektedir. Kodlayıcıların içerik hakkında güvenilir kararlar alabilmeleri için, kesin tanımlar kullanılması ve titiz bir eğitimden geçmiş olmaları zorunludur. Örneğin, ergenlerin kendini yaralama davranışlarına ilişkin bir çalışmada araştırmacılar, ergenin kendini yaralama davranışıyla ilgili internet forumlarının içeriğini analiz etmek için ikili bir kod dizisi (var/yok) kullandılar (Whitlock, Powers ve Eckenrode, 2006). Bu kodları, internette ilan edilmiş mesajları gözleyerek ve kendini yaralayanlarla görüşerek elde ettiler. Daha sonra, iki aydan fazla bir süre, 10 internet forumundan 3219 internet ilanı üzerinde çalıştılar ve kendini yaralama motivasyonu ve bu davranışı gizleme yolları gibi mesajın içeriğindeki farklı temalara göre gruplandırdılar. Veri indirgeme tekniği kullanılarak kodlama yapılması, araştırmacıların belirli tür davranışlarla bu davranışlardan önce meydana gelen olaylar arasındaki ilişkileri belirlemesine izin verir. Whitlock ve arkadaşları yaptıkları kodlamada örneğin, kendini yaralama davranışının "tetikleyicilerini" tespit ettiler ve her bir tetikleyiciyi açıklayan mesaj oranını tespit edebildiler. Bu kodlamalarına dayanarak, kendini yaralamanın en sık tekrarlanan tetikleyicisinin (%34.8) "önemli bulunan diğer insanlarla çatışma" olduğunu gözlediler.

Biz bu kitapta niceliksel veri analizine ağırlık verdik. Bu nedenle, niteliksel veri analizine sadece özet bir giriş yapma olanağı bulduk. Psikolojik araştırmaların bu önemli yaklaşımı hakkında daha fazla bilgi edinmek isteyen öğrenciler, niteliksel yöntemin pek çok sistematik açıklamasına ulaşabilirler (örneğin, Miles, Huberman ve Saldana, 2013; Saldana, 2013; Silverman, 2011).

Niceliksel Veri Analizi

- Veriler sıklık, ortalama, standart sapma gibi betimsel istatistikler kullanılarak özetlenir.
- Gözlemciler arası güvenirlik, bağımsız gözlemcilerin aynı gözlemde ne ölçüde uyuştuklarını ifade eder.
- Gözlemciler arası güvenirlik, kayıt edilecek olay ve davranışlar için açık ve net tanımlar yapılarak, gözlemciler eğitilerek ve gözlemlerin doğruluğu hakkında geri bildirim sağlanarak arttırılır.
- Yüksek bir gözlemciler arası güvenirlik, araştırmacıların davranış hakkındaki gözlemlerinin doğru (geçerli) olması olasılığını arttırır.
- Gözlemciler arası güvenirlik, davranışların nasıl ölçüldüğüne ve kaydedildiğine bağlı olarak, uyuşma yüzdesi veya korelasyon hesaplanarak değerlendirilir.

*Niceliksel veri analizi*nin amacı, çalışmadaki gözlemlerin sayısal veya niceliksel bir özetini ortaya koymaktır. Önemli bir aşama, gözlem verilerini özetleyen frekans, ortalama ve standart sapma gibi betimsel istatistikleri hesaplamaktır. Gözlem verilerinin analizinin bir diğer önemli boyutu, gözlemlerin güvenirliğini değerlendirmektir. Gözlemler güvenilir olmazsa, davranışlar hakkında anlamlı bir şey söylememiz mümkün değildir. Şimdi niceliksel veri analizinin bu yönlerinin her birisini, sırayla açıklayacağız.

Betimsel İstatistikler Gözlem verilerini özetlemekte kullanılacak betimsel istatistiğin türü, veriyi kaydetmekte kullanılan ölçek türüne bağlıdır. Daha önce gördüğümüz gibi kategorik ölçekler, davranışlar ve olaylar ayrı kategoriler altında toplandığı zaman kullanılır. Sıklıkla kullanılan kategorik ölçüm bir davranışın varlığı ya da yokluğu olduğundan, kategorik ölçekler için en yaygın betimsel istatistik, frekanstır (sıklık – frequency). Frekansı hesaplamak için davranış veya olayın kaç defa ortaya çıktığı sayılır ve sonra gözlemlerin toplam sayısına bölünür. Frekans ölçümü oran veya yüzde olarak ifade edilir. Daha önce de sözünü ettiğimiz gibi, Whitlock ve meslektaşları, ergenlerde kendini yaralama davranışının tetikleyicilerini kodlamışlardı ve tekrarlanma sıklığı en yüksek tetikleyici "önemli insanlarla çatışma" idi. Tetikleyicilerden söz edilen 609 mesaj içinde çatışma konusu geçenlerin 212 tane olduğunu sayarak belirlediler. O hâlde frekans, .348'dir (212/609) veya mesajların %34.8'i çatışma tetikleyicisini içermektedir.

Sıralama verilerini tanımlarken araştırmacılar, genellikle madde dizisi içinde birinci sırada olan en yüksek sıklıktaki maddeyi bildirirler. Vatandaşların ülkeleri hakkındaki düşüncelerine yönelik bir ankette araştırmacılar, katılımcılardan hükümetin öncelik vermesi gerektiğini düşündükleri ekonomi, savaş, eğitim, çevre, ulusal güvenlik vb. sorunlar hakkında sıralanmış maddeleri derecelendirmelerini isteyebilir. Sonuçları rapor edecekleri zaman insanların birinci sırada derecelendirdiği maddenin yüzdesini bildirebilirler; "katılımcıların % 35'i hükümetin ekonomiye öncelik vermesi gerektiğini düşünüyor" gibi (hipotetik veri). Daha eksiksiz bir tanımlama, birinci sıradaki yüzdeden sonra geriye kalanları da içerir; katılımcıların % 28'i çevre sorunlarına öncelik verilmesi gerektiğini düşünüyor; %.. gibi. Sıralama verilerini tanımlamanın bir diğer yolu, madde dizisindeki seçilmiş belirli maddeler için, 1., 2., 3. gibi, dereceler ile yüzdeleri birlikte vermektir. Hipotetik verimize göre bunu şöyle ifade edebiliriz: "Katılımcıların %35'i ekonomiyi birinci sırada önemli buluyor, %25'i ekonomiyi ikinci sırada, %12'si üçüncü sırada ... önemli buluyor" gibi.

Farklı –ve çok daha bilgilendirici– bir betimsel istatistik, davranış en azından eşit aralıklı ölçeklerle kaydedildiğinde rapor edilebilir. Gözlemler eşit aralıklı ölçek veya oranlı ölçek (süre, gecikme) kullanılarak kaydedildiği zaman, bir veya daha fazla merkezî eğilim ölçüsü kullanılır. En yaygın kullanılan merkezi eğilimi ölçüsü aritmetik ortalamadır. Ortalama, bir grubun puanlarının "tipik" puanını tanımlar ve bir birey veya grubun performansını özetlemek için kullanışlı bir ölçüm sağlar. Performansın daha eksiksiz bir tanımlaması için, araştırmacılar puanların ortalama etrafındaki dağılımını veya bir ölçümdeki değişkenlikleri de bildirmelidir. *Standart Sapma*, bir puanın ortalamaya yaklaşık olarak ne kadar uzak olduğunu belirtir.



Şimdi, verileri sistematik şekilde analiz etmeye yarayacak genel esasları olduğu kadar, merkezî eğilim ölçülerini ve değişkenliği incelemek için de uygun bir aşamaya geldiğimizi söyleyebiliriz.

Betimsel istatistiklerin kullanımını bir örnekle açıklamak için aşağıdaki çalışmayı inceleyelim. Pek çok insan, farklı kültürlerden gelen bireylerin temel kişilik özellikleri bakımından farklı olduklarına inanır. Örneğin, Meksikalılar sıklıkla Amerikalılardan daha dışa dönük ve sosyal görünür. Bununla beraber, dışa dönüklüğü öz bildirim ölçekleriyle belirleyen kültürler arası çalışmalar, genellikle Meksikalıların kendilerini Amerikalılardan daha az dışa dönük bulduklarını gösteriyor (Ramirez-Esparza, Mehl, Alvarez-Bermudez ve Pennebaker, 2009). Yaygın kanıya uygun olmayan bu bulguyu nasıl açıklayacağız? Bir olasılık, bireylerin nasıl davrandıklarına ilişkin algılarının (öz bildirim) aslında nasıl davrandıklarıyla örtüşmemesidir. Bu açıklamayı sorgulamak için araştırmacılar doğal ortamlardaki sosyal davranışları sistematik olarak gözlediler (Ramirez-Esparza ve ark., 2009).

Çalışmalarında, katılımcıların bellerine veya cüzdan benzeri bir çantaya taktıkları elektronik olarak faaliyete geçen cihaz (EHG) kullandılar (Bu bölümdeki Kutu 4.1'e bakınız). EHG periyodik olarak, katılımcının bulunduğu ortamdaki sesleri kaydedip ileterek, katılımcıların günlük hayatlarına işitsel bir pencere açmış oldu. Amerikalı ($N = 46$) ve Meksikalı ($N = 46$) üniversite öğrencilerinden hafta içi iki gün mümkün olduğu kadar çok EHG takmaları istendi. EHG, 12.5 dakikada bir 30 saniye kayıt almaya programlanmıştı. Araştırma asistanları bu kayıtları 4 kategori kullanarak kodladılar: Yer (örneğin, ev, ulaşım aracı gibi), etkinlik (örneğin, yemek yemek, müzik dinlemek gibi), etkileşim (örneğin, yalnız, konuşurken), duygu durumu (örneğin, gülmek, ağlamak gibi). Bu dört kategori yoluyla "davranışsal sosyallik ortak ölçümü" elde edildi [ayrıntılar için Ramirez-Esparza ve arkadaşlarına (2009) bakınız].

Tablo 4.4'te EHG'ten sağlanan sosyallik puanlarının ortalama ve standart sapmaları verilmiştir. Ortalamalar, Meksikalı üniversite öğrencilerinin Amerikalı öğrencilerden daha sosyal davrandığını ve her iki milliyette de erkeklerin kadınlardan daha sosyal olduğunu göstermektedir. Standart sapma değerleri, Amerikaların Meksikalılardan daha az değişkenlik gösterdiğine işaret etmektedir. Diğer yandan, standart kâğıt-kalem öz bildirim ölçeklerinde Amerikalılar kendilerini Meksikalılardan daha sosyal gördüklerini ifade etmişlerdi. Bu da bize, öz bildirim ölçeklerinin gerçek davranışla örtüşmeyebileceğini göstermektedir.

TABLO 4.4 AMERİKALI VE MEKSİKALI ÖĞRENCİLERİN SOSYALLİK DÜZEYİNİ BETİMLEYEN ORTALAMA VE STANDAR SAPMA DEĞERLERİ*

Grup	Ortalama	Standart Sapma
Amerikalılar		
Erkekler ($n = = 25$)	25.37	10.46
Kadınlar ($n = = 29$)	20.28	9.18
Toplam	22.64	10.03
Meksikalılar		
Erkekler ($n = = 20$)	36.51	16.79
Kadınlar ($n = = 26$)	30.79	16.27
Toplam	33.27	13.63

*Nairan Ramirez-Esparza'dan alınmıştır.

Gözlemci Güvenirliği Betimsel istatistiklere ek olarak araştırmacılar, çalışmadaki her gözlemin ne kadar güvenilir olduğunu araştırırlar. Güvenirliğin tutarlılığa dayandığını hatırlayacaksınız; gözleme dayalı araştırmaların güvenirlik analizinde, bağımsız gözlemcilerin aynı olayları incelediklerinde benzer sonuçlara ulaşıp ulaşmadıklarına bakılır. İki (veya daha fazla) bağımsız gözlemcinin uyuşma derecesi **gözlemciler arası güvenirlik** olarak tanımlanır. Eğer gözlemciler uyuşmamışsa, gerçekte meydana gelmiş ve ölçülmüş olan olay veya davranış ile ölçülmüş olanın örtüştüğünden emin olamayız. Gözlemciler arası güvenirliğin düşük olmasının nedeni, gözlenmiş olayların açıkça tanımlanmamış olması olabilir; ve/veya gözlemcilerin davranışları kaydederken öznel davranmış olmaları olabilir. Gözlemciler arası güvenirliği arttırmak için araştırmacılar, çok net sözel tanımlamalar yapmanın yanında gözlenecek davranışın video kayıtları veya fotoğraflarını içeren somut örnekler verebilir. Gözlemciler arası güvenirlik genellikle gözlemcilerin eğitilmesi ve gözlem yapma becerilerini geliştirme fırsatları sağlanarak geliştirilir. Gözlemcilere eğitim verilirken ve gözlem pratikleri yaptırılırken, birbirlerinin gözlemleri arasındaki tutarsızlıklara ilişkin geri bildirim vermek özellikle yararlıdır (Judd, Smith ve Kidder, 1991).

Temel Kavram

İki gözlemcinin gözlemleri uyuştuğu zaman, genellikle tek bir gözlemcinin gözlemine dayanan bir veriden daha geçerli ve isabetli bir gözlem yapıldığına olan inancımız artar. Gözlemcilerin bağımsız olabilmesi için her biri, bir diğerinin kayıt yaptığının farkında olmamalıdır. Her iki gözlemcinin aynı zamanda ve aynı derecede beklentiler, yorgunluk veya bıkkınlıktan etkilenmiş olması olasılığı düşüktür. Böylece biz, onların bildirimlerinin gerçekte olmuş olanın ne olduğu hakkında bir uyuşma sağladığından emin olabiliriz. Şüphesiz daha fazla bağımsız gözlemcinin uyuşması daha emin olmamız anlamına gelir.

Gözlemciler arası güvenirliğin değerlendirilme biçimi, davranışın nasıl ölçüldüğüne bağlıdır. Olaylar bağımsız kategorilere (kategorik ölçeğe) göre sınıflanmışsa, gözlemci güvenirliği uyuşma yüzdesi kullanılarak değerlendirilir. Gözlemciler arası uyuşma yüzdesi aşağıdaki formüle göre hesaplanır:

$$\frac{\text{İki gözlemcinin uyuşma sayısı}}{\text{Uyuşma fırsatlarının sayısı}} \times 100$$

Düşük gözlemciler arası güvenirliği belirlemek için kesin bir kural yoktur ama araştırmacılar basılmış yayınlarda genellikle daha düşük uyuşma yüzdesi kabul edilmediği izlenimini bırakacak şekilde %85'i aşan güvenirlik aralıkları rapor ederler.

Veriler, sıralama ölçeğiyle ölçülmüşse, gözlemciler arası güvenirliği değerlendirmek için Spearman sıra-farkları korelasyonu kullanılır. Gözleme dayalı veri eşit aralıklı veya oranlı ölçeklerle ölçülmüşse Pearson Çarpım Momentleri korelasyon katsayısı (r) ile değerlendirilir. Facebook profillerinden sağlanan bilgiler kullanılarak kişilerin kişilik özelliklerinin değerlendirildiği araştırma, bunun bir örneğiydi. Örneğin, kodlayıcılar, ilgi sayısı gibi nesnel bir ölçüt kullanarak Facebook kullanıcılarının profillerini inceledikten sonra, onların sosyal

İSTANBUL TİCARET ÜNİVERSİTESİ
KÜTÜPHANESİ

endişe düzeylerini derecelemek için bir ölçek kullandılar (Fernandez ve ark., 2012). Korelasyonlar, gözlemciler arası güvenirliğin göstergesi olarak kodlayıcıların sosyal kaygı düzeyi için yaptıkları derecelemeler arasında uyuşma olduğunu gösteriyordu.

Temel Kavram

İstatistik İpucu

Aynı insan, olay veya şeylerden alınan iki farklı ölçme birlikte değişiyorsa, aralarında *korelasyon* var demektir. Yani bir değişkendeki puanlarla öteki değişkendeki puanlar birlikte değişmektedir. **Korelasyon katsayısı**, bu birlikte değişimin derecesinin sayısal bir göstergesidir. Gözleme dayalı bir veri aralıklı veya oranlı bir ölçek kullanılarak ölçüldüğü zaman, Pearson korelasyon katsayısı, r, gözlemciler arası güvenirlik ölçümünü elde etmek için kullanılabilir. Korelasyon bize iki gözlemcinin uyuşma derecesini bildirir.

Korelasyon katsayısı, ilişkinin hem yönünü hem de gücünü gösterir. Yön pozitif veya negatif olabilir. Pozitif korelasyon, bir değişken arttıkça diğerinin de arttığını gösterir. Örneğin, sigara içme ile akciğer kanseri ölçümleri pozitif yönde ilişkilidir. Negatif korelasyon, bir değişken arttığı zaman diğerinin azaldığını gösterir. Örneğin, televizyon izlemeye harcanan zaman ile akademik test puanları negatif yönde ilişkilidir. Gözlemciler arası güvenirlik değerlendirildiği zaman araştırmacılar pozitif korelasyon bekler.

Korelasyonun gücü, belirlenen birlikte değişimin düzeyini ifade eder. Korelasyonun yayılım genişliği -1.00 (mükemmel negatif ilişki) ile 1.00 arasındadır. 0 değeri iki değişken arasında ilişki olmadığını gösterir. Korelasyon katsayısı 1.00 veya -1.00'e yaklaştıkça, iki değişken arasındaki ilişki güçlenir. Dikkat edilmelidir ki korelasyonun işareti yalnızca yönünü ifade eder. Bir korelasyon katsayısı -.46 ise, .20'den güçlü demektir. Gözlemciler arası güvenirlik ölçümlerinde, .85'i aşan değerlerin gözlemcilerin yüksek düzeyde uyuştuğunu gösterdiğini söylemiştik (fakat daha yükseği, daha iyidir!).

Bölüm 5'te, korelasyondan yararlanılarak kestirimlerde bulunmayı tartışacağız. Bölüm 11'de ayrıca, korelasyonun ayrıntılı bir tartışması, iki değişken arasındaki ilişkinin dağılım grafi ği kullanılarak nasıl grafi kle gösterileceğini, Pearson Çarpım Momentleri Korelasyon katsayısının nasıl hesaplanacağını ve korelasyon değerlerinin en iyi nasıl yorumlanacağını da ele alacağız.

GÖZLEME DAYALI ARAŞTIRMALAR ÜZERİNE ELEŞTİREL DÜŞÜNME

İyi bir gözleme dayalı araştırma, davranış ve olay örneklerinin nasıl gözleneceğini, uygun gözlem yönteminin nasıl seçileceğini ve gözlem verilerinin nasıl kayıt ve analiz edileceğini belirler. Şimdi temel gözleme dayalı yöntemleri bildiğinize göre, oluşabilecek potansiyel sorunları da bilmeniz gerekir. Birinci sorun, gözlemci gözleyeceği davranışı etkilediği zaman ortaya çıkar;

ikinci sorun ise gözlemci, kaydedeceği davranışı yanlı olarak seçtiği zaman ortaya çıkar. Bu iki sorunu sırasıyla ele alacağız.

Gözlemcinin Etkisi

- Tepkisellik sorunu, bir gözlemcinin varlığı, gözlenen davranışı etkilediği zaman ortaya çıkar.
- Araştırmaya katılanlar, araştırma ortamında "kendilerinden beklenen" şekilde davranıyor olabilirler.
- Tepkiselliği kontrol yöntemleri şunlardır: Gözlemcinin varlığının gizlenmesi, uyum sağlama (alışma, duyarsızlaşma) ve dolaylı gözlem (fiziksel izler, arşiv kayıtları).
- Araştırmacılar, tepkiselliği kontrol etmek istedikleri zaman etik meseleleri de göz önüne almalıdır.

Tepkisellik Gözlendiklerini bilme, gözlenen insanların davranışlarını değiştirmelerine yol açabilir. Katılımcı gözlemleri açıkladığımız bölümde, bu tepkisellik meselesi üzerinde durmuştuk. Bireyler, bir gözlemcinin varlığına "tepki" verdikleri zaman, davranışları, tipik davranışlarını temsil etmeyebilir. Katılımcılar davranışlarının gözlendiğini fark ettikleri zaman, çok incelikli yollarla tepki verebilirler. Bazen bir psikolojik araştırmaya katıldıkları için endişeli olabilirler ve uyarılmışlık ölçümleri (örneğin, kalp atışı) gözlemcinin varlığı yüzünden değişebilir. Elektronik bir uyarıcı takan katılımcılara, davranışlarının ve duygu durumlarının kaydedileceği sinyali verildiği zaman, beklendiği gibi, davranışlarını değiştirdiler (örneğin, Larson, 1989). Tepkisellik sorunu EHG yöntemi kullanan araştırmacıların da sorunudur (örneğin, Mehl ve Holleran, 2007). Elektronik bir kayıt cihazıyla donatılınca, katılımcılar belirli aralıklarla konuşmalarının kayıt edileceğinin farkına varırlar. Beklenebileceği üzere, bazı katılımcılar, önemli bazı kimselerle mahrem konuşmalar yaparken cihazı kapattıklarını itiraf ettiler.

Bireyler çoğunlukla, gözlemcinin varlığına, araştırmacının beklentisine uygun olduğunu düşündükleri şekilde davranarak tepki verirler. Bilimsel bir araştırmanın parçası olduğunu bilen bireyler genellikle işbirliği yapmak ve "iyi" bir katılımcı olmak isterler. Katılımcılar sıklıkla kendilerinden ne beklendiğini tahmin etmeye çalışırlar ve bu tahminlerini kolaylaştıracak ipuçlarını kullanabilirler (Orne, 1962). Araştırma ortamındaki bu ipuçlarına **talep özellikleri** denilir. Örne, bireylerin kendilerine genellikle "Burada bulunmam ne işe yarıyor?" sorusunu sorduklarını düşünüyor. Bu soruyu cevaplarken katılımcılar mevcut durumdaki ipuçlarına, araştırmanın işlemlerine ve araştırmacı tarafından verilen örtük ipuçlarına dikkat ediyorlar.

Temel Kavram

Katılımcı talep özelliklerine dikkat ederek davranışını değiştirdikçe, araştırmanın dış geçerlilik olasılığı da düşer. Katılımcılar, araştırma ortamı dışındaki davranışlarından farklı bir şekilde davrandıkları zaman, araştırma bulgularının genellenebilirlik yeteneği (dış geçerlik) zayıflamış olur. Bu durum ayrıca araştırma bulgularının doğru şekilde yorumlanmasını da tehdit eder çünkü katılımcılar farkında olmadan, bağımsız değişkeni olduğundan daha etkili hâle getirebilir veya bir başka önemli değişkenin etkilerini değiştirebilirler. Talep özelliklerini

azaltmanın tek yolu, mümkün olduğu kadar az "ipucu" vererek katılımcılardan araştırmada ne beklendiğini ve çalışmanın hipotezi hakkındaki bilgileri sınırlandırmaktır. Bununla birlikte hatırlayacağınız üzere katılımcıdan bilgi saklamak, özellikle bilgilendirilmiş onam düşünüldüğünde etik sorunlar oluşturabilir.

Tepkiselliğin Kontrol Edilmesi Tepkisellik sorununu kontrol etmek için araştırmacıların kullandığı bir kaç yaklaşım vardır. Bu bölümün başında ele aldığımız bir kaç gözlem yöntemi, tepkiselliği sınırlandıracak şekilde tasarlanmıştır. Tepkisellik, eğer katılımcılar gözlemcinin ortamda mevcut olduğunu bilmiyorlarsa ortadan kalkar. Gizlenmiş katılımcı gözlem bu amaca ulaşılmasını sağlar, çünkü bireyler bir gözlemcinin mevcudiyetinin farkında değildir. Gözlemciler doğal ortamda gözlem yaparken (doğal gözlem) kendilerini saklamış olurlar veya gözlemlerinde gizli kamera veya ses kayıt cihazı kullanabilirler (fakat mahremiyet ihlali ile ilişkili etik meselelerin farkında olmaları gerekir).

Dolaylı gözlem veya dikkat çekmeyen yöntemlerin önemli bir üstünlüğü bu gözlemlerde tepkiselliğin olmamasıdır. Araştırmacılar insanların geçmiş davranışlarını öğrenmek için fiziksel izler veya arşiv kayıtlarını inceler. Bireyler ortamda olmadıkları ve bir araştırmacının kendilerinden kalan fiziksel izleri veya arşivleri gözleyeceklerini bilmedikleri için, davranışlarını değiştirmeleri imkânsızdır. Bir araştırmacı, resmî olarak içkinin "yasak" olduğu bir mahallede yaşayan insanların çöp kutularındaki boş içki şişelerini sayarak onların içki içme davranışlarını araştırdı (Şekil 4.5'e bakınız). Bir diğer araştırmacı bir topluluğa televizyonun girişinin etkilerini değerlendirmek için kütüphanenin arşiv kayıtlarını kullandı. Televizyonun gelmesinden sonra kütüphaneden alınan bilim kurgu kitaplarının sayısında düşüş olmuştu ama diğer kitapların sayısında fark olmamıştı [Webb ve arkadaşlarına (1981) bakınız].

Araştırmacıların tepkisellikle baş etmekte kullandıkları diğer bir yaklaşım katılımcıların, gözlemcinin ortamdaki varlığına alışmalarıdır. Katılımcılar gözlemcinin mevcudiyetini kanıksarsa, gözlemci varken de normal şekilde davranacaklarını varsayabiliriz. Uyum sağlama, duyarsızlaştırma veya alıştırma ile gerçekleştirilebilir. Bir *alıştırma* prosedüründe gözlemci, sadece katılımcılar onun varlığına alışsın diye, artık ona tepki vermeyene kadar araştırma ortamına defalarca girer çıkar (yani, araştırmacının ortamdaki varlığı artık normal hâle gelir). EHG yöntemi kullanan araştırmacılar, tepkiselliği azaltmak için alışmaya bel bağlıyorlar (örneğin, Mehl ve Holleran, 2007). Araştırmacılar şuna işaret ediyorlar:

> EHG takıldıktan hemen sonra, kısa bir dönem katılımcılar EHG'in sıklığı üzerine konuşmalarında, abartılmış bir farkındalık dönemi geçiriyorlar. Ama cihazın takılışından sonraki 2 saat içerisinde, pek çok katılımcı bu yönteme alışıyor ve daha sonra diğer insanlara nadiren bundan söz ediyor (s. 254).

Tepkiselliğin muhtemel etkileri devam etmekle birlikte, araştırmacılara göre, "potansiyel olarak diğer insanların gözledikleri günlük olaylara odaklanıldığı zaman", tepkiselliğin etkileri en aza inecektir. (s. 256). Muhtemeldir ki benzeri bir alışma süreci günümüzdeki "realite şovlar"da da görülmektedir, fakat bu şovlarda gösterilen bazı davranışların tam olarak öyle (yani, göründüğü gibi) olup olmadığı da ayrı bir sorundur, çünkü en nihayet insanlar televizyondadır!

ŞEKİL 4.5 İnsan davranışlarının dikkat çekmeyen (tepkisel olmayan) ölçümleri, çöp kutularındaki fiziksel izleri inceleyerek de elde edilebilir. Ancak mahremiyet gibi etik sorunlara dikkat edilmelidir.



Tepkisellikle baş etmenin bir yolu olarak *duyarsızlaştırma* (desensitization), davranışçı fobi tedavilerinde klinik psikologlar tarafından kullanılan yönteme benzer. Bir tedavi ortamında, belirli bir korkusu (örneğin, örümcek) olan birisi, önce o korkuyu yaratan uyarana çok düşük bir düzeyde maruz bırakılır. Örneğin bireyin örümcek ağı gibi, örümcekle ilgili şeyler üzerine düşünmesi istenir. Eş zamanlı olarak terapist, hastanın kas gevşetme alıştırması yapmasına yardım eder. Hasta, korktuğu şeyin gerçek hâline dayanabilir (örneğin örümceği tutabilir) hâle gelene kadar aşamalı bir şekilde uyaranın yoğunluğu arttırılır. Duyarsızlaştırma, sıklıkla hayvan araştırmacıları tarafından gözlemcinin mevcudiyetine katılımcı hayvanların uyum sağlaması amacıyla kullanılır. Afrika'da vahşice öldürülmeden önce Dian Fossey (1981, 1983) Ruanda'daki dağ gorilleri üzerinde hayranlık uyandıran gözleme dayalı araştırmalar yürütmüştü. Çok uzun bir zaman süresince, goriller onun varlığına uyum sağlayana kadar, giderek daha fazla yakınlaşacak şekilde ilerledi. Onların hareketlerini taklit etmenin (mesela onların yediği yeşillikleri çiğnemek ve kendini kaşımak) gorilleri rahatlattığını keşfetti. Nihayetinde kendisine dokunacakları ve araştırma araç-gereçlerini inceleyecekleri kadar gorillere yakın olmayı başardı ve ancak bu aşamadan sonra onları gözledi.

Etik Meseleler Araştırmacılar, insanların haberi olmadan onları gözleyerek tepkiselliği kontrol ettiklerinde önemli etik meseleler ortaya çıkar; örneğin, ciddi bir mahremiyet (özel hayat) ihlali olabilir. Nelerin özel hayatın ihlali sayılacağına karar vermek kolay değildir (Bölüm 3'e bakınız) ve gözlemin yapıldığı ortam,elde edilen bilginin hassasiyeti, bunların ne şekilde yayılacağı göz önüne alınmalıdır (örneğin, Diener ve Crandall, 1978).

EHG yönteminde, mikrofonun yakınında oldukları için konuşmalarının kaydedildiğinin farkında olmayan başka insanlar olabilir; bu durum bir mahremiyet sorunu teşkil eder ve araştırmacılar en baştan bu gibi sorunların farkındadırlar. (Mehl ve ark., 2001, s.522). Bu hem etik hem yasal sorunlara neden olmaktadır. Bununla beraber, katılımcı olmayanların seslerini dinleyerek, çoğunlukla kim oldukları ortaya çıkarılamaz; bu durumda o insanların özel hayatlarının korunduğu düşünülebilir. Ama yine de bu olasılık, araştırmacıların EHG'yi yeniden düzenlemesine yol açmıştır; örneğin, mikrofon yalnızca katılımcının sesini alacak şekilde giysinin altına yerleştirilmektedir (Mehl ve ark., 2010). Ayrıca, konuşma kayıtları yazıya dökülmeden önce, konuşmalar dinletilmekte ve onların uygun bulmadıkları konuşmalar çıkarılmaktadır.

İnternet kullanan yeni davranışsal araştırmalar yeni bir etik ikilemle karşılaştılar. Örneğin, araştırmacılar, ırkçıların ırkçı şiddeti savunmak için neler yaptıklarını öğrenmek için internet sohbet odalarına gizlenmiş gözlemci olarak katıldılar (Glaser ve ark., 2002); ama öğrendikleri şeyler, polisin bir "suçüstü" operasyonu yapmasını gerektirecek önemde bilgilerdi ve katılımcıların bundan haberi yoktu. Bu durumdaki ikilem şudur: Eğer katılımcılardan bilgilendirilmiş onam alınmış olsaydı bu bilgilerin alınması olanaksızdı. Bu vakada etik kurul, sohbet odalarının bir "kamusal alan" (umumî) olduğu, bu konuların bu forumlarda yaygın olduğu ve araştırmacıların, katılımcıların kimliğini korumak için yeterli önlemleri aldıkları sonucuna vardı ve araştırmayı onayladı. Öte yandan, insanların kendilerinden izin alınmadan internetteki konuşmalarını araştırmacının dinlemesini, mahremiyet ihlali olarak kabul ettikleri örnekler vardır (Skitka ve Sargis, 2005'e bakınız). İnternet forumları "kamusal alan" olarak kabul edilebilmekle birlikte, üniversite etik kurulları, ergenlerin kendilerini yaralama davranışlarına ilişkin internet mesajlarını inceleyen araştırmacıların, bu mesajları aynen yayımlamak yerine, mesajları özetleyerek farklı şekilde yazmalarını istedi (Whitlock ve ark., 2006). İnterneti kullanan davranış araştırmaları, bu yöntemin henüz başlangıç aşamasındadır; dolayısıyla, hem araştırmacılar hem de etik kurul üyeleri, etik ikilemleri çözmek için yaratıcı yöntemleri öğrenmeye devam etmektedir (Buchanan ve William, 2010; Kraut ve ark., 2004).

Bireyler araştırmacının amaçlı şekilde düzenlediği bir ortama karıştırıldıkları zaman, yapılandırılmış gözlemler veya alan deneylerinde olduğu gibi, katılımcının üstlendiği risk ile bağlantılı etik sorunlar doğar. Örneğin bir alan deneyinde kampüste yürüyen öğrenciler, ırkçı taciz hakkındaki soruları cevapladılar (Blanchard, Crandall, Brigham ve Vaughn, 1994). İki deney koşulundan birisinde öğrenci rolüne girmiş bir işbirlikçi ırkçı eylemleri kınadı, ikinci koşuldaki işbirlikçi ise hoş gördü. Sonra katılımcılara bireysel olarak tutumları hakkında sorular soruldu. Araştırmanın sonuçları gösterdi ki işbirlikçinin açıkladığı bir görüşe maruz kalan katılımcılar, bir işbirlikçinin görüş belirtmediği kontrol koşuluna kıyasla, işbirlikçilerin ifadelerine daha fazla benzeyen görüşler bildirdiler. Bu katılımcıların hangi açıdan "risk altında" olduklarını söyleyebiliriz? Farklı ırklardan insanların yaşadıkları ortamlarda, açık sözlü insanların etkisini göstermek isteyen bu araştırmanın amacı, muhtemel risklerinden daha önemli miydi? Bu araştırmada katılımcılar hemen bilgilendirilmiştir ama bu bilgilendirme, şöyle bir endişeyi gidermeye yeter mi? Acaba katılımcılar ırkçı görüşlerle gerçekten karşılaşsaydılar, öyle mi davranırlardı? Yine aynı bilgilendirme, bilgi elde etmek için aldatma

yöntemini kullanan bir bilime karşı sarsılan güveni onarmaya yeter mi? Bu soruları bu gibi konularda etik kararlar almanın güçlüklerini göstermektedir.

Son olarak, fiziksel izler veya arşiv kayıtları gibi dikkat çekmeyen ölçümlerdeki bir diğer etik meseleye dönebiliriz: Bilim insanlarının, bireysel ve toplumsal koşulları geliştirme konusundaki etik yükümlülüğüne. Bugün yüzleşmemiz gereken şiddet, ırk ilişkileri, intihar, aile içi çatışmalar ve bunlar gibi pek çok ciddi sorun vardır. Bu nedenle, doğrudan gözlemi kullanan araştırmaların fayda/yarar oranları göz önüne alındığı zaman, bunları savunmak güç olabilir. Yani bazı araştırma yöntemleri, araştırmaların katılımcıları için büyük riskler içerebilir. Arşiv verileri ve fiziksel izler yöntemlerini kullanan araştırmalar, bu önemli sorunları azaltıyor olabilirler. Bu nedenle dikkat çekmeyen gözleme dayalı yöntemler, önemli sosyal sorunların araştırılması için çok yöntemli araştırmaların önemli bir seçeneğidir.

Gözlemci Yanlılığı

- Gözlemci yanlılığı, araştırmacıların gözlemek için hangi davranışları seçeceklerini belirlerken ortaya çıkabileceği gibi, gözlemcinin davranış hakkındaki beklentileri davranışın tespitinde ve kaydedilmesinde sistematik hatalara yol açtığı zaman da ortaya çıkar.
- Beklenti etkisi, gözlemciler araştırmanın hipotezlerinin veya önceki araştırmaların sonuçlarının farkında oldukları zaman ortaya çıkar.
- Gözlemci yanlılığını kontrol etmenin birinci aşaması, yanlılık olabileceğinin farkında olmaktır.
- Gözlemci yanlılığı, gözlemcilerin araştırmanın hipotezleri ve amaçlarından habersiz (kör) olmalarını sağlayarak azaltılabilir.

Gizlenmiş katılımcı gözlemin bir örneği olarak, Rosenhan'ın (1973) gözlemcilerin bir psikiyatri hastanesine hastaymış gibi başvurup yatırıldıkları klasik araştırmasından söz etmiştik. Onlar hastanede ilk olarak hastane çalışanlarının davranışlarını gözlediler ve kaydettiler. Rosenhan'ın araştırmasında çalışanların bir kısmında ciddi yanlılıklar tespit edilmişti. Bir kere gözlemcilere (sahte hastalara) şizofreni teşhisi konulduktan sonra, çalışanlar onların davranışlarını yalnızca bu etikete göre yorumladılar. Bu sahte hastalara ne teşhis konulduğunu bilmeselerdi "normal" olarak niteleyebilecekleri birçok davranışı, onlara konulan teşhise göre yorumladılar. Örneğin, sahte hastalar gözlemlerini alenen kaydedebildiklerini çabucak fark ettiler, çünkü hiç kimse onların ne yaptığıyla ilgilenmiyordu; ne yapsalar "normal" karşılanıyordu, çünkü onlar "hasta"idi. Rosenhan sonradan bu sahte hastalar hakkındaki tıbbî kayıtları incelediği zaman, çalışanların bu not alma davranışını onların hastalığının bir belirtisi olarak kaydedip bildirdiklerini buldu (Endişe etmeyin; not almak, bir zihinsel rahatsızlık belirtisi değildir!). Çünkü çalışanlar sahte hastaların davranışlarını, şizofreniye göre yorumluyordu; bunların "sağlıklı" davranışlar oldukları keşfedilemedi. Bu örnek **gözlemci yanlılığının** tehlikelerini açıkça göstermektedir: Gözlemcinin beklentilerinin bir sonucu olarak, gözlemdeki sistematik hatalar. Bu olayda hastane çalışanları, gözlemci yanlılığı göstermiştir.

Temel Kavram

Beklenti Etkisi Pek çok bilimsel araştırmada gözlemciler, belirli bir psikolojik müdahale veya belirli bir durumda hangi davranışların ortaya çıkacağı hakkında bazı beklentilere sahiptir. Araştırmacılar bir çalışma tasarlarken daha önceki araştırmalardan yola çıkarak hipotezlerini kurarlar. Bu bilgi,

araştırmacıların araştırma ortamında neyin ortaya çıkması gerektiği hakkında bir beklenti geliştirmelerine yol açabilir; aslında bir hipotez, ortaya çıkması beklenen davranış hakkındaki bir öngörüdür. Bununla birlikte, beklentiler ancak gözlemde sistematik hatalara yol açıyorsa, bir gözlemci yanlılığı –*beklenti etkisi*– kaynağı sayılabilir (Rosenhan, 1966, 1976). Beklenti etkisi, klasik bir araştırmada gösterilmiştir (Cordaro ve Ison, 1963). Üniversite öğrencileri, yassı kurtların beden kasılmalarının ve kafa hareketlerinin sayısını gözleyerek kaydettiler. Bir grup gözlemcide yüksek hareket beklentisi, diğer grupta ise düşük hareket beklentisi oluşturuldu. Gerçekte ise her iki grup yassı solucan aynıydı. Fakat sonuçlar gösterdi ki fazla hareket görmeyi bekleyen grup, az hareket görmeyi bekleyen gruba oranla aynı sayıda kafa hareketini 2 kez ve aynı sayıda beden kasılmasını üç kez kaydetmişti. Görünüşe göre, öğrenciler ne görmeyi beklemişlerse, solucanların hareketlerini sistematik şekilde ona göre yorumlamışlardı.

Diğer Yanlılıklar Gözlemci yanlılığının tek kaynağı, bir gözlemcinin çalışmanın sonucuna ilişkin beklentisi değildir. Açıktır ki video kameraları gibi kendisi çalışan donanımlar kullanılarak, gözlemci yanlılığı ortadan kaldırılabilir. Ama otomasyon yanlılık şansını azaltmakla birlikte, bütünüyle ortadan kaldırmaz. Şöyle düşünün, film kaydı için araştırmacılar bir açı, bir yer ve zaman belirlemek zorundadırlar. Araştırmanın bu boyutları bir ölçüde araştırmacının kişisel yanlılığından etkilenmektedir; böyle kararlar, sonuçlara sistematik hatalar getirebilir. Ayrıca, otomatik donanımların kullanılması genellikle yalnızca sınıflama ve yorumlama sürecini ertelemiş olur; kayıtlar kodlandığı ve analiz edildiği zaman, gözlemci yanlılığı etkisinin işe karışması pekâlâ mümkündür.

Gözlemci Yanlılığının Kontrol Edilmesi Gözlemci yanlılığının ortadan kaldırılması zordur ama bir kaç yol izlenerek azaltılabilir. Daha önce belirttiğimiz gibi, otomatik kayıt donanımlarının kullanılması yardımcı olabilir fakat yanlılık potansiyel olarak hâlâ mevcuttur. *Muhtemelen gözlemci yanlılığıyla baş etmek için en önemli faktör böyle bir yanlılığın ortaya çıkabileceğinin farkında olmaktır.* Yani, yanlı olabileceğinin farkında olan bir gözlemcinin, bu etkiyi azaltmak için önlem alması olasılığı daha yüksektir. Araştırmacıların gözlemci yanlılığını azaltmanın diğer önemli bir yolu, gözlemcilere verilen bilginin sınırlandırılmasıdır. Gözlemciler ve kodlayıcılar çalışmanın hipotezini bilmezlerse, davranışlar hakkında bir beklenti geliştiremezler. Bir başka deyişle gözlemciler, çalışmanın belirli yönleri hakkında "kör" kalabilirler. Gözlemciler kör oldukları zaman çalışmanın amacını veya gözlemin nedenlerini bilmezler. Örneğin tacize uğramış veya uğramamış ailelerden gelen anne-çocuk etkileşiminin video kayıtlarını analiz edecek kodlayıcılar eğitilirken hangi tür aileleri gözlediklerini bilmezler (Valentino ve ark. 2006). Tahmin edebileceğiniz gibi, gözlemcinin tacize uğramış ailelerle ilgili beklentileri, davranışları yorumlamasını etkileyebilir. Kör gözlemciler kullanılarak, gözlemci beklentisi nedeniyle sistematik hataların ortaya çıkması olasılığı büyük ölçüde azaltılabilir.

ÖZET

Araştırmacılar nadiren bütün davranışları gözleyebilir. En sonunda araştırmacıları yapmaları gereken şey, ortamdan örnek seçme ve zamandan örnek seçme gibi bazı davranış örnekleri seçmektir. Örneklemenin en önemli amacı, "temsil edici" bir davranış örneklemine ulaşmaktır. Dış geçerlik, çalışmadaki gözlemlerin ne ölçüde farklı evren, ortam ve koşullara genellenebileceği demektir. Temsil edici bir örneklem sağlandığı zaman, dış geçerlik artar. Gözleme dayalı yöntemler doğrudan gözlem ve dolaylı gözlem olarak sınıflanabilir. Bir müdahale yapılmaksızın doğal ortamda doğrudan gözlem yapmaya, doğal gözlem denilir. Müdahale yapılan gözlemler katılımcı gözlem, yapılandırılmış gözlem, ve alan deneyidir. Dolaylı gözlem yöntemlerinin en önemli üstünlüğü tepkisel olmamalarıdır. Tepkisellik, katılımcıların, gözlendiklerini bildikleri için davranışlarını değiştirdikleri zaman ortaya çıkar. Dolaylı veya farkına varılmayan gözlemler, fiziksel izler ve arşiv kayıtlarının incelenmesiyle elde edilir. Fiziksel izler, kullanım izlerini (doğal veya kontrollü) ve ürünleri kapsar. Arşiv verileri bireyler, kurumlar, devlet veya diğer grupların etkinliklerinin kaydıdır. Fiziksel izlerle ilgili sorunlar, izlerin zaman boyunca nasıl biriktirildiği ve korunduğuna ilişkin potansiyel yanlılığı kapsar; arşiv verilerinin sorunları seçici birikim, seçici korunma ve verideki sahte ilişki olasılıklarıdır.

Gözleme dayalı çalışmalarda, davranışın kapsamlı bir betimlemesi yapılabilir, veya davranışın yalnızca önceden belirlenmiş belli yönleri kaydedilebilir. Anlatı kayıtları davranışın kapsamlı bir betimlemesini sağlar; kontrol listeleri ise, araştırmacılar belirli bir davranışın meydana gelip gelmediğiyle (ve meydana gelmişse, hangi koşullar altında meydana geldiğiyle) ilgilendiği zaman kullanılır. Sıklık, süre ve davranışın oranı, gözleme dayalı araştırmalarda yaygın olarak incelenen değişkenlerdir. Anlatıların içerik analizi ve arşiv kayıtları, bir veri indirgeme aşamasını gerektirir. Niceliksel verilerin nasıl analiz edileceği, hangi ölçme düzeyinin kullanıldığına bağlıdır. Ölçme düzeyleri nominal, sıralama, eşitlik aralıklı ve oranlı olmak üzere dört tiptir. Davranışı kaydetmek için nominal düzey (örneğin, var-yok) kullanıldığı zaman veriler, davranışın görece sıklığını gösteren yüzde veya oran kullanılarak özetlenir. Sıralama verisi kullanıldığı zaman araştırmacılar sıklıkla bir dizi maddenin ilk sırasındakileri işaretleyen katılımcıların yüzdesine göre sonuçları betimler. Eşit aralıklı veya oranlı ölçekler kullanıldığı zaman veriler ortalama ve standart sapma ile özetlenir. Gözleme dayalı araştırmaların sonuçları rapor edildiği zaman gözlemci güvenirliği ölçümlerinin yapılması gerekir. Kullanılan ölçüm düzeyine bağlı olarak ölçümlerin uyuşma yüzdesi veya korelasyon katsayısı güvenirliği değerlendirmek için kullanılabilir.

Gözleme dayalı araştırmalarda, tepkisellik veya gözlemci yanlılığından kaynaklanan olası sorunlar kontrol edilmek zorundadır. Tepkiselliğin bir şekli, katılımcılar araştırma ortamında kendilerinden beklendiğini düşündükleri şeyleri (talep özellikleri) yapmaya çalıştıkları durumda ortaya çıkar. Katılımcının gözlendiğinin farkında olmadığı (örneğin, gizlenmiş katılımcı gözlem, fark edilmeyen gözlem gibi) gözleme dayalı yöntemler, tepkisellik olasılığını azaltır; aynı sonuç, katılımcılar gözlemcinin varlığına alıştıkları zaman da görülür. Gözlemci yanlılığı, araştırmacılar araştırmacılar gözlemek istedikleri davranışları seçerken

yanlı davrandıkları zaman ortaya çıkabilir; ayrıca, yine araştırmacılar davranışı tespit ederken ve kaydederken sistematik hataya yol açan beklentilere (beklenti etkisi) sahip oldukları zaman da ortaya çıkar. Gözlemci yanlılığının azaltılmasında önemli bir adım, gözlemcinin, yanlılık olabileceğinin farkında olmasıdır; bir diğer önlem ise çalışmanın amacını ve hipotezini bilmeyen kör gözlemciler kullanmaktır. Etik meseleler, her gözleme dayalı araştırmaya başlamadan önce mutlaka ele alınmalıdır. Gözlemlerin niteliğine bağlı olarak, etik meseleler kandırma, mahremiyet, bilgilendirilmiş onam ve risk/fayda oranını içerir.

TEMEL KAVRAMLAR

dış geçerlik 94
zaman örnekleme 94
durum örnekleme 96
doğal gözlem 99
katılımcı gözlem 101
tepkisellik 101
yapılandırılmış gözlem 103
alan deneyi 105
dikkat çekmeyen ölçümler 106
fiziksel izler 107
arşiv kayıtları 110
seçici birikim 111
seçici kalıcılık 112
anlatı kayıtları 113
ölçüm düzeyi 115
içerik analizi 119
kodlama 119
veri indirgeme 119
gözlemciler arası güvenirlik 123
korelasyon katsayısı 124
talep özellikleri 125
gözlemci yanlılığı 129

GÖZDEN GEÇİRME SORULARI

1 Gözleme dayalı araştırmalarda, araştırmacıların kullandığı örnekleme şekillerini ve iyi bir örnekleme yapılması için nelerin gerektiğini açıklayın.
2 Doğrudan ve dolaylı gözlem arasındaki farkları ve müdahale düzeyinin doğrudan gözlem yöntemlerini ayırt etmek için nasıl kullanılabileceğini açıklayınız.
3 Araştırmacının davranışa müdahale etmesini engelleyen etik endişeler olduğu zaman kullanılabilecek doğal gözleme dayalı bir araştırma durumu betimleyin.
4 Gözleme dayalı araştırmalarda tepkiselliğin niçin bir sorun olduğunu açıklayın.
5 Bir katılımcı gözlem araştırması internette yürütüldüğü zaman hangi etik sorunlara ortaya çıkabilir?
6 Fiziksel izlerin ve arşiv verilerinin niçin doğrudan gözlemlerin cazip bir alternatifi olduğunu açıklayın.
7 Psikologlar için uygun fiziksel iz ölçümlerinin farklı türlerini ve bu ölçümlerin her birinin nasıl yanlı olabileceğini açıklayın.
8 Arşiv verileri doğal bir müdahalenin etkisini sınamak üzere kullanılabilir mi? Açıklayın.
9 Seçici birikme, seçici kalma ve sahte ilişkiler, arşiv verilerinin yorumlanmasına nasıl yanlılık getirebilir? Açıklayın.
10 Araştırmacıların birbiriyle konuşan çiftler arasında göz teması kurma davranışını nasıl ölçebileceğine, dört ölçek türünün her birisini kullanarak örnek veriniz.
11 Arşiv kayıtları ve anlatı kayıtlarının niteliksel analizinde veri indirgeme ve kodlamanın nasıl kullanıldığını açıklayın.
12 a) Nominal ölçek kullanıldığı zaman; b) sıralama ölçeği kullanılarak maddeler derecelendirildiği zaman ve c) davranış en az aralıklı ölçekle kaydedildiği zaman yararlanılan en yaygın betimsel ölçümleri açıklayın.
13 Araştırmacıların gözlemciler arası güvenirliği arttırmak için kullanabilecekleri işlemleri açıklayın.

14 Gözlemciler arası güvenirliği değerlendirirken korelasyon katsayısı kullanmayı gerektiren ölçme düzeylerini belirleyin ve böyle bir durumda negatif korelasyonun ne anlama geldiğini açıklayın.
15 Psikolojik araştırmalarda ortaya çıkan gözlemci yanlılığının (beklenti etkisi) iki türünü açıklayın.
16 Araştırmacıların gözlemci yanlılığını nasıl azaltabileceklerini açıklayın.

DÜŞÜNDÜRÜCÜ SORULAR

1 Öğrenciler gelişim psikolojisi laboratuvar dersinde evdeki aile-çocuk etkileşimi üzerine gözleme dayalı bir araştırma yürütmektedirler. Belirlenen bir aileyi 4 günde bir gözleyecekleri eve ilk kez girdikleri zaman, hem ebeveynlerle hem de küçük çocukla tanıştılar (diğer çocuklar da evdeydi). Aileye, günlük olağan işlerine devam etmelerini söylediler ve o günün, aile için olağan bir gün olup olmadığını anlamak amacıyla o günkü etkinlikler hakkında sorular sordular. Öğrenciler ailenin kendisini rahat hissetmesini sağlamaya çalıştılar ama birbirleriyle ve aileyle etkileşimlerini en az düzeyde tutmaya çalıştılar. İki saatlik gözlem süresinde her zaman 2 öğrenci evde bulunuyordu ve bunlar, gözlemlerini birbirlerinden bağımsız olarak kaydediyorlardı. Altı öğrenci çifti, çalışmaya katılmaya gönüllü olan 12 aileye yansız olarak atanmıştı. Aynı gözlemci çifti, aynı aileyi toplam 8 saat süreyle gözlüyordu. Gözlemciler açık şekilde belirlenmiş bir ölçek kullandılar, ve aile-çocuk etkileşimini sıcaklık ve yakınlık gibi çok sayıda davranış boyutuna göre kaydettiler.
   - **A** Bu bulguların güvenirliğini sağlamak için öğrencilerin kullandıkları iki belirgin işlemin hangileri olduğunu söyleyin.
   - **B** Bu araştırmanın bulgularının dış geçerliliği için muhtemel bir tehdit söyleyin ve betimlenen ortamdan spesifik bir örnek verin.
   - **C** Bu ölçümde tepkisellik oluşturması mümkün olan ve öğrencilerin tepkisellik olabileceğinin farkında olduklarını gösteren bir işlem söyleyin. Tepkisellik sorunuyla başa çıkmak için kullanılabilecek başka yöntemler hangileridir?

2 Gözleme dayalı bir araştırmada, üniversitenin sponsor olduğu bir birahanede, çevresel faktörlerin üniversite öğrencilerinin içme davranışı üzerindeki etkileri değerlendirildi. Yirmi bir yaşın üzerinde 82 öğrenci gözlendi. Gözlemciler bir kontrol listesi kullanarak katılımcıların cinsiyetini, bir kişiyle mi yoksa iki veya daha fazla kişiden oluşan bir grupla mı birlikte olduğunu kaydettiler. Her gözlem oturumu, daima saat 15.00 ile 01.00 arasındaydı ve gözlemler pazartesiden cumartesiye kadar yapıldı. Gözlemler 3 aydan fazla sürdü. Her gözlem oturumunda iki gözlemci vardı. Her katılımcı, ilk birasını aldıktan sonraki 1 saat boyunca gözlendi. Veriler her bir saatte içilen bira sayısı cinsinden özetlendi. Sonuçlar gösterdi ki erkekler kadınlardan daha çok ve daha hızlı içiyor. Erkekler bir erkekle birlikteyken daha hızlı içiyordu ve kadınlar da bir erkekle birlikteyken daha hızlı içiyordu. Hem kadınlar hem de erkekler, bir grupla birlikteyken bir tek kişiyle birlikte olduklan durumdan daha hızlı içiyordu. Bu sonuçlar, içki içilen çevrenin içmenin düzeyi ve niteliği üzerinde önemli bir rol oynadığına işaret etmektedir.
   - **A** Bu çalışmada kullanılan gözleme dayalı yöntemi açıklayın ve neden bu yöntem olduğuna karar verdiğinizi açıklayın.
   - **B** Bu çalışmadaki bağımlı ve bağımsız değişkenleri belirtin ve bağımsız değişkenin her düzeyinin işevuruk tanımını yapın.
   - **C** Araştırmacılar bu çalışmadaki tepkiselliği nasıl kontrol edebilirdi? Yaklaşımlarından ne gibi etik sorunlar doğabilirdi?
   - **D** Bu çalışmada gözlemlerin güvenirliğini arttırması muhtemel olan işlemlerin bir yönünü tanımlayın.
   - **E** Bu çalışmada, bulguların dış geçerliliğini sınırlaması muhtemel olan işlemlerden birisini açıklayın.

3 On altı yaşına kadar Almanya'da büyümüş bir Amerikalı psikoloji öğrencisi, Almanlar ve Amerikalıların sosyal paylaşım sitelerini nasıl kullandıklarını araştırmak istiyor. Akıcı şekilde Almanca konuşmaktadır ve hem Almanya'da hem de Amerika'da sosyal bağlantıları vardır. Öğrenci özellikle şunu anlamak istiyor: Amerikalıların ve Almanların Facebookta kendilerini tarif etme biçimleri arasında bir fark var mıdır? Her iki ülkedeki arkadaşlarıyla bağlantı kurarak projesini açıklıyor. Ayrıca, bunu kendi ülkelerindeki arkadaşlarıyla da paylaşmalarını istiyor. Bir kaç hafta sonra her iki ülkeden yaklaşık olarak aynı sayıda, 500'den fazla facebook kullanıcısı katılımcı buluyor. Siz bu öğrenci araştırmacıya, içerik analiziyle ilgili aşağıdaki sorular çerçevesinde ne gibi önerilerde bulunursunuz?

(devam ediyor)

A Niceliksel bir analiz, niteliksel bir analizden ne şekilde farklıdır?
B Niceliksel analizin bir parçası olarak ne kodlanabilir?
C Kodlamanın güvenirliği nasıl değerlendirilebilir?
D Öğrencinin örnekleminin hangi özellikleri, sonuçlarının dış geçerliğini etkileyebilir?

4 Dört öğrenci üniversitelerinin Sosyal Bilim Araştırmaları Enstitüsünde staj yapıyorlar. Araştırma enstitüsü, üniversiteye yakın küçük bir şehrin şehir merkezini geliştirme ajansıyla, bir dizi trafik güvenliği araştırması yapmak üzere anlaştı. Stajyer öğrenciler şehir merkezindeki yayaların yürüdüğü ve dur işareti olan kavşaklarda, araçların durmaya ne kadar eğilimli olduklarını belirleyeceklerdi. Öğrencilerin araştırmalarını planlarken göz önüne aldıkları hususlarla ilgili aşağıdaki soruları cevaplayın.

A Öğrenciler araçların eğilimini, "evet - "hayır" şeklinde sınıflamaktan daha detaylı şekilde derecelemek istiyor. Öğrenciler araçların durması, yavaşlaması ve hiç durmaması durumlarını ayırt edebilecek bir işevuruk tanımı nasıl geliştirebilirler?
B Öğrenciler esas çalışma için veri toplamaya başlamadan önce, gözlemlerinde gözlemciler arası güvenirliği arttırmak için hangi adımları atabilirler?
C Öğrenciler yaya trafiği fazla ve az olduğu durumlarda da araçların durma eğilimlerini belirlemek istiyor Öğrenciler bunu belirlerken, hangi zaman-örnekleme planını kullanmalıdır?
D Öğrenciler özellikle diğer durmadıkları hâlde, durma işaretinde duran araçlarla ilgilenmektedir. Bu şekilde duran araçları diğer araçlardan bağımsız olarak incelemek için nasıl bir örnekleme yapmaları gerekir? Öğrenciler, örneklemlerindeki bütün araçları kaydetmenin yanı sıra, bağımsız olarak duran araşları da kaydetmek için hangi bilgileri kaydedebilir?

İleri Alıştırmaların Cevapları

1 Doğal gözlem kullanılmıştır. Davranış müdahale yapılmaksızın doğal ortamda gözlenmiştir.
2 Örnekleme işlemleri yaşantı örnekleme (yemekler) ve sistematik belirlenmiş aralıklarla yapılan zaman örneklemenin bir birleşimi olarak betimlenebilir (5 günden fazla bir süre, bir saatte dört defa 15 dakikalık aralarla).
3 Yalnızca bir kişi elektronik kayıt cihazıyla gözlendiğinde, erkeklerin oranı (40/50 =.80), kadınların oranından (10/50= .20) büyüktür. Bu sonuç öğrencilerin görüşünü desteklemektedir.
4 Gözlemcilerin, gözleyecekleri çiftleri nasıl seçtikleri açık değildir. Örneğin, onlar elektronik cihazları masanın üzerinde duran çiftlere daha çok dikkat etmiş olabilirler. Gözlemcilerin basitçe kimi isterlerse seçmelerinden ziyade, gözleyecekleri çiftleri seçkisiz olarak belirlemeleri daha uygundur. Ayrıca, araştırmacılar müdahalenin süresi ya da şekliyle ilgili veri toplamamışlardır. Bir aramayı cevaplamak ve çabucak kapatmak, 10 dakika Angry Bird oynamaktan daha farklı kodlanmış olabilir. Üstelik, araştırma asistanlarının aynı çiftleri gözleyerek ve karşılaştırarak, gözlemciler arası uyum ölçümü yapılabilirdi.

Düşündürücü Soruların Cevapları

A Öğrencilerin güvenirliği artırma işlemleri şöyledir: Her ailenin 8 saat gözlenmesi; 2 bağımsız gözlemci kullanılması; davranışın işevuruk tanımını veren açıkça tanımlanmış bir ölçek kullanılması.
B Bulguların dış geçerliğine olası bir tehdit, 12 ailenin de gönüllü olması ve böyle ailelerin tipik ailelerden farklı olabilecekleridir.
C Öğrenciler, tepkisellik sorununa duyarlı oldukları için aileyle ve birbirleriyle etkileşimi azaltma çabasına girdiler. Sekiz saatlik bir gözlemde aile üyeleri gözlemcinin mevcudiyetine alışabilir. Bir diğer yöntem, gözlemcilerin giderek daha fazla süre aile içinde bulunarak aileleri duyarsızlaştırma olabilir.

BÖLÜM BEŞ

Anket Araştırmaları

BÖLÜMÜN ANA HATLARI

GENEL BAKIŞ
ANKETLERİN KULLANIMI
ANKET ARAŞTIRMALARININ ÖZELLİKLERİ
ANKET ARAŞTIRMALARINDA ÖRNEKLEM SEÇME
Örnek Seçmenin Temel Kavramları
Örnekleme Seçme Yaklaşımları
ANKET YÖNTEMLERİ
Posta Anketleri
Yüz Yüze Görüşmeler
Telefon Görüşmeleri
İnternet Anketleri
ANKET ARAŞTIRMA DESENLERİ
Kesitsel Desen
Ardışık Bağımsız Örneklemler Deseni
Boylamsal Desen
ANKET FORMU
Bir Araç Olarak Anket Formları
Öz Bildirim Ölçümlerinin Güvenirlik ve Geçerlikleri
Bir Anket Formunun Hazırlanması
ANKET ARAŞTIRMALARI ÜZERİNE ELEŞTİREL DÜŞÜNME
Bildirilen Davranış ile Gerçek Davranış Arasındaki Uyum
İlişkisellik (Korelasyon) ve Nedensellik
ÖZET

GENEL BAKIŞ

Amerikalılar romantik midir? Peki ya aşka aşık olan Fransızlarla karşılaştırıldıklarında romantik midirler? Bu sorular, 2009'da Amerikan romantizmi üzerine yapılan ve sonuçları aşk ve aşk ilişkileri konusunda bir Fransız araştırmasının sonuçlarıyla karşılaştırmak isteyen bir ankette sorulan sorulardan iki örnektir.

Anket sonuçları, Amerikalıların da Fransızlar kadar "aşık" olduğunu gösteriyor. Hatta daha yaşlı katılımcılar söz konusu olduğunda Amerikalılar, Fransızlardan daha "aşık". Altmış beş yaşın üzerindeki Amerikalıların % 63'ü kendini "aşık" olarak tanımlarken aynı yaş grubundaki Fransızlarda bu oran % 46. Bu iki grup, cinsellik söz konusu olduğunda da farklılaşıyor. "Tutkulu bir seks hayatı olmadan gerçek aşk mümkün olabilir mi?" sorusuna 18-65 üstü yaştaki Amerikalı katılımcıların çoğunluğu (% 77) "evet" derken, Fransızların sadece % 35'i tutkulu bir seks yaşamı olmadan da gerçek aşkın yaşanabileceğini ifade etmiş.

Bu sonuçlara dayanarak, insanların "aşık olmak" ile ilgili tepkilerini betimleyebiliriz. Ayrıca, yaşa ve uyruğa (Amerikalı ya da Fransız) göre "aşık olmak" ile ilgili tepkileri de tahmin edebiliriz. Elde edilen bulgular ayrıca, katılımcının Amerikalı ya da Fransız olmasına göre gerçek aşk ve seks ile ilgili ne söyleyebileceğini tahmin etmemize de olanak verir. Ancak Fransız ya da Amerikalı olmak bu tutumlara neden olur mu? Bu tamamen farklı bir konudur.

Temel Kavram

Korelasyonel araştırmalar, tahminlerde bulunmak için bir temel teşkil eder. Değişkenler arasında doğal olarak ortaya çıkan bağlantılar, *yordayıcı ilişkileri* belirlemek amacıyla değerlendirilir. Bölüm 4'te tartıştığımız gibi *korelasyon katsayısı*, yordayıcı ilişkinin yönünü ve gücünü gösteren sayısal bir değerdir. Anket yöntemleri bağlamında ilişkisel araştırmaları, bu bölümde daha sonra ele alacağız.

Anketler genellikle insan örneklemleriyle yürütülür. Bu bölümde ilk olarak örneklem seçimindeki (örnekleme) temel mantık ve teknikler üzerinde duracağız. Örneklem seçimi, bir evrenin (popülasyonun) tamamını temsil eden bir alt kümesini seçme sürecidir. Farklı araştırma yöntem ve desenlerinin avantaj ve dezavantajlarını göreceğiz. Bir anketin temel aracı bir soru listesidir. Bu nedenle iyi bir soru formu oluşturmak için yapılması gerekenleri açıklayacağız. Aynı zamanda bir araştırmada cevaplanması gereken önemli bir soruyu tartışacağız: "İnsanlar gerçekten söyledikleri şeyi yaparlar mı?" Bölümü, "İki değişken arasında bir korelasyon olduğunda, nedensellikle ilgili bir yargıda bulunabilir miyiz?" sorusunu eleştirel bir şekilde ele alarak bitireceğiz.

ANKETLERİN KULLANIMI

- Anketler, insanların düşünce, görüş ve duygularını değerlendirmek için kullanılır.
- Anketler amaçlarına göre özelleşmiş ve dar kapsamlı ya da daha genel olabilir.
- Bir anket sonucunun yanlı olup olmadığını anlamanın en iyi yolu, araştırmanın işlem sürecinin ve analizlerinin incelenmesidir.

Bölüm 4'te psikologların, insanların belirli bir şekilde davranmak için ne düşündüklerini ya da ne hissettiklerini anlamak için gözleme dayalı yöntemleri nasıl kullandıklarını ele almıştık. Anketler insanların duygu, düşünce ve görüşlerine daha doğrudan ulaşmak için tasarlanır. Yüzeysel olarak

değerlendirildiğinde, anket araştırmaları yanıltıcı şekilde basit görünür: Ne düşündüklerini bilmek istiyorsan, kendilerine sor! Aynı şekilde, ne yaptıklarını bilmek istiyorsan, insanları gözle. Bununla birlikte daha önce de gördüğümüz gibi, davranışın genel ilkelerini keşfetmeyi amaç edinmişsek, gözlemlerimiz, gündelik gözlemlerimizden çok daha incelikli olmalıdır. Bu nedenle anket araştırmaları, insanlara basitçe soru sormaktan fazlasını yapmayı da gerektirir.

Siyaset bilimciler, psikologlar ve sosyologlar gibi sosyal bilimciler, araştırmalarında hem teorik ve hem de uygulama açısından farklı nedenlerle anket araştırmalarını kullanır. Anketler medyanın, siyasi adayların, kamu sağlığı yetkililerinin, şirketlerin ve satış-pazarlama yöneticilerinin pratik ihtiyaçlarını karşılamak için kullanıldığı gibi, filmlerde insanların sigara içerken gösterilmesini engellemek isteyen bir kamu sağlığı hareketinde olduğu gibi, politik ya da toplumsal amaçları desteklemek amacıyla da kullanılır. Heatherton ve Sargent'in (2009) analiz ettiği anket verilerine göre, filmlerde sigara içildiğini görmek özellikle ergenlerin sigarayı deneme olasılıklarını ya da sigaraya başlama olasılıklarını artırıyor; özellikle de ebeveynleri sigara içmeyen düşük risk grubundaki ergenlerin.

Anket araştırmalarının kapsamı ve amacı sınırlı ve özelleşmiş olabileceği gibi küresel de olabilir. Dar kapsamlı bir araştırmada (Lambert ve ark., 2010), bireylerin yakın ilişkilerinde minnet duygularını ifade etme derecesi, ilişkinin gücü hakkındaki algılarıyla korelasyonlu bulunmuştur. Daha geniş kapsamlı bir araştırmada, çok sayıda katılımcıyla oynanan bir bilgisayar oyunundaki bireylerin tutumlarını incelemek için, hem online görüşmeler hem de e-posta yazışmaları kullanılmıştır (MMORPGs; bk., Hussain & Griffiths, 2009). Niteliksel olan bu araştırma için 11 farklı ülkedeki online oyunculardan veri toplanmıştır. Küresel mutlulukla ilgili bilgi toplamayı amaçlayan daha geniş kapsamlı anketler de mevcuttur (örneğin, Ingelhart, Foa, Peterson ve Wetzel, 2008).

Anket araştırmalarının bazı kullanımları bazı etik kaygılara neden olmaktadır. Araştırmanın sponsorları araştırma sonuçlarından çıkar sağlayacaksa, etik bir ikilem ortaya çıkmaktadır. Crossen (1994) bu durumun önemini şöyle vurguluyor: "Bir ürünü satın almak, önermek, seçmek ya da aklamak için kullandığımız araştırma sonuçları, bilgimizi artırmaktan çok, o ürünün satışını ya da satın alma nedenini desteklemek için üretilmektedir" (s, 14). Crossen, cep telefonu üreticilerinin sponsorluğunda yapılan bir araştırmadan örnek veriyor: Hepsi birer telefon kullanıcısı olan katılımcıların %70'i, cep telefonu kullanan kişilerin kullanmayanlara oranla iş yaşamında daha başarılı olduğu konusunda hemfikir!

Peki, sonuçları sponsorların çıkarına hizmet eden araştırma bulgularının her zaman yanlı olduğunu düşünmek mantıklı mıdır? Etik soruların cevapları nadiren basittir ve yukarıdaki sorunun basit bir cevabı yoktur. Sonuçları sponsorun çıkarlarına hizmet etse de kaliteli ve etik bir araştırma yürütmek mümkündür. Bulguları değerlendirirken sponsoru biliyor olmak önemlidir ancak bu durum, araştırmanın yanlı olduğunu söylemek için yeterli değildir. Araştırmada kullanılan örneklemin yanlı olup olmadığını, sorularda kullanılan ifadelerin yönlendirici olup olmadığını ya da verilerin bir kısmının seçilerek analiz ve rapor edilip edilmediğini bilmek çok daha önemlidir. Bu faktörlerin her birisi, sonuçların yanlı olmasına neden olabilir ve etik davranmayan araştırmacılar sonuçları "düzeltmek" için bu teknikleri kullanabilirler. Etik davranmayan araştırmacılardan ve düşük kaliteli araştırmalardan korunmanın en iyi yolu, araştırmada kullanılan işlem ve analizleri incelemektir.

ANKET ARAŞTIRMALARININ ÖZELLİKLERİ

- Anket araştırmaları bir örneklem (ya da örneklemler) seçmeyi ve önceden hazırlanmış bir soru dizisi kullanmayı gerektirir.

Kurallara uygun olarak yürütülen bütün anketler, insanların tutum ve davranışlarını belirlemek için kullanılan bu yöntemi geliştiren bazı ortak özelliklere sahiptir. Bu özelliklerin ilki, neredeyse tüm davranışsal araştırmaların özelliği olan örneklem seçmedir. Bu kavrama, Bölüm 4'te yer alan gözleme dayalı araştırmalar kısmında değinmiştik. Anketlerde kullanılan örneklem seçme tekniklerini ileride ele alacağız. Anketlerin ikinci özelliği, bütün katılımcıların cevaplayacağı önceden hazırlanmış soru listelerinin kullanılmasıdır. Bu sorulara sözlü, yazılı ya da bilgisayar ortamında verilen cevaplar, anketle elde edilen esas verileri oluşturur. Aynı ifade ve soru sıralamasını kullanarak, tüm katılımcıların görüşlerini kısaca özetlemek mümkündür.

Temsil edici bir örnek gruptaki kişilere aynı sorular sorulduğu zaman, örneklemin alındığı evrenin tutumlarını belirleyebiliriz. Hatta aynı sorular kullanılarak farklı evrenlerin tutumlarını karşılaştırabilir ya da zamanla tutumlarda meydana gelen değişiklikleri belirleyebiliriz. Anketler, araştırmacıların alet çantasındaki güçlü bir araçtır. Bu bölümün sonunda anketleri insanların duygu, düşünce ve görüşlerini araştırmak için etkili bir strateji haline getirecek yöntemlere dikkat çekeceğiz.

ANKET ARAŞTIRMALARINDA ÖRNEKLEM SEÇME

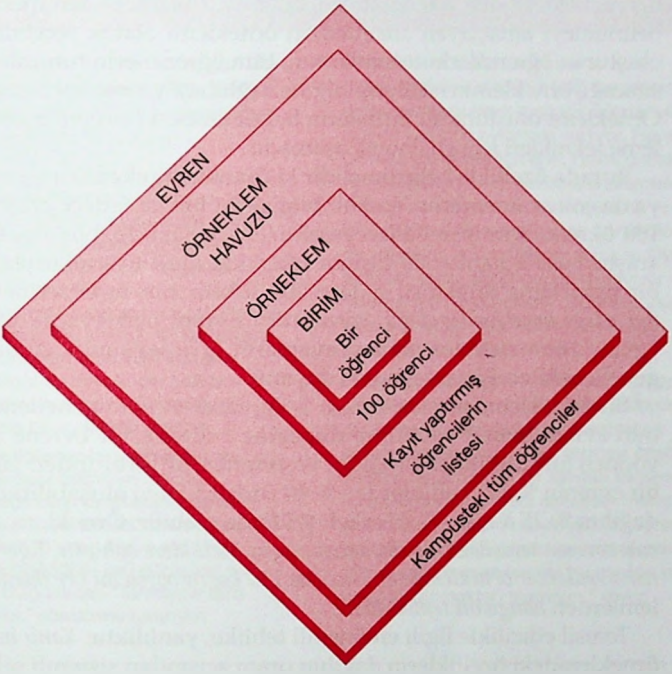
- Uygun tekniklerle seçilmiş bir örneklem, araştırmacıların bu örneklemden elde edilen bulguları evrene genellemesini mümkün kılar.

Diyelim ki araştırma sorunuza en iyi şekilde cevap almak için bir anket kullanmaya karar verdiniz ve anketinizle ilgili bir evren belirlediniz. Sonraki adım, araştırma sorularınıza kimlerin cevap vermesi gerektiğine karar vermektir. Bunun için evreni temsil edecek katılımcı örnekleminin dikkatlice seçilmesi gerekir. İster ulusal bir evren ister daha küçük bir evren (örneğin bir üniversitenin öğrencileri) belirlensin, temsil ediciliği olan bir örnek seçmek için kullanılacak işlemler aynıdır.

Örnek Seçmenin Temel Kavramları

- Örneklemi oluşturacak birimlerin tanımlanması ve seçimi, tüm örnekleme tekniklerinin merkezinde yer alır. Örneklem, ya örnekleme havuzundan ya da araştırmaya konu olan evrenin tüm üyeleri arasından seçilir.
- Araştırmacılar, sadece araştırmaya katılanların cevaplarıyla ilgilenmezler. Bunun yerine, örneklemin içinden alındığı daha büyük evreni açıklamak isterler.
- Örneklemden elde edilen bulguların evrene genellenebilmesi, örneklemin temsil ediciliğine bağlıdır.
- Yanlı bir örneklem, evrenin özelliklerinden sistematik olarak farklı özelliklere sahip olan örneklemdir.

ŞEKİL 5.1 Örnek Seçmeyle ilgili dört temel kavram arasındaki ilişkilerin şeması



Eğer örneklem seçmede kullanılan teknik evrenin bazı kısımlarını aşırı ya da eksik temsil ederse, yanlı seçim ortaya çıkar.

Temel Kavram

Örneklem seçme tekniklerinden söz ederken kullandığımız dört kavramın –*evren, örnekleme havuzu, örneklem, birim*– tanımları konusunda açık olmamız gerekir. Bu dört kavram arasındaki ilişkiler Şekil 5.1'de özetlenmiştir. Bir **evren**, hakkında veri toplanmak istenilen tüm birimlerin toplamıdır. Örneğin kampüsünüzdeki öğrencilerin bilgisayar servisleriyle ilgili tutumlarını öğrenmek istiyorsanız, evreniniz kampüsteki tüm öğrencilerdir. Böylesine büyük bir evrendeki herkese ulaşmak genellikle çok zordur. Bu nedenle araştırmacılar çoğunlukla, evrenin içinden o evrenin tamamını temsil eden bir alt küme seçerler.

Evrene ait bir alt küme seçebilmek için o evreni oluşturan tüm üyelerin bulunduğu belirli bir liste oluşturmamız gerekir. Bu belirli listeye örnekleme havuzu denir ve bu kavram bir anlamda, ilgilenilen evrenin işevuruk (operasyonel) tanımıdır. Öğrencilerin bilgisayar servisleriyle ilgili tutumlarının inceleneceği araştırmada örneklem havuzu, kayıt yaptırmış tüm öğrencilerin bulunduğu bir liste olabilir. Seçtiğimiz örneklemin uygunluğu, örnekleme havuzunun evreni ne kadar temsil ettiğine (onunla aynı özellikleri ne kadar taşıdığına) bağlıdır. Kayıt ofisinden aldığımız bu liste iyi bir örnekleme havuzu sağlar, ancak geç kaydolan bazı öğrenciler bu havuzun dışında kalmış olabilir.

Temel Kavram

Örnekleme havuzundan seçilen evrenin alt kümesine, **örneklem** denir. Kayıtlı öğrenci listesinden 100 öğrenciyi, bilgisayar hizmetinden memnuniyet düzeyini belirlemeyi amaçlayan anketimizin örneklemi olarak seçebiliriz. Bu örneklemi oluşturan öğrencilerin tutumlarının, tüm öğrencilerin tutumlarını ne kadar yansıtacağı, örneklemin nasıl seçildiğine bağlıdır. Evrenin her bir üyesine *birim* denir. Örneklemi oluşturacak birimlerin belirlenmesi ve seçimi aşaması, bütün örnekleme teknikleri için en önemli aşamadır.

Burada özellikle belirtilmelidir ki örneklemin kendisi ya hiç önemli değildir ya da çok az önemlidir. Yeni bir bilgisayar binası sadece araştırmaya katılan bu 100 öğrenci için inşa edilmeyecektir. Aynı şekilde, bir sosyal psikolog sadece araştırmasına katılan 50 kişinin ırkçılıkla ilgili tutumlarıyla ilgilenmez ya da bir pazarlama direktörü sadece 200 tüketicinin tercihlerini önemsemez. *Esas ilgi odağı örneklemler değil, evrenlerdir.* Evreni açıklayacak olan örneklemlerin "gücü", örneklemden alınan cevapların, o örneklemin içinden alındığı evrene genellenebileceği varsayımına dayanır.

Temel Kavram

Bir örneklemden elde edilen bulguların evrene genellenebilmesi, örneklemin evreni **temsil edebilme düzeyine** bağlıdır. Bir evrene ait bireyler birçok yönden birbirlerinden farklıdır, ve evrenler de birbirinden farklı olur. Örneğin, bir evrenin % 60'ı kadınlardan % 40'ı erkeklerden oluşabilirken, bir diğerindeki dağılım % 25 kadın, % 75 erkek şeklinde olabilir. Örneklem, genelleme yapılacak *evrenin temsilcisi olarak, evrenle aynı özelliklere sahiptir. Eğer 200 kişiden oluşan temsil edici bir örneklemde 80 erkek ve 120 kadın varsa bu örneklem yukarıdaki* örneklemlerden *hangisini temsil eder?*

Temel Kavram

Temsil edicilikle ilgili en önemli tehlike, yanlılıktır. *Yanlı bir örneklem* demek, örneklemdeki özelliklerin dağılım oranı açısından sistemli şekilde hedef evrenden farklı olması demektir. 80 kadın ve 20 erkek, toplam 100 kişiden oluşan bir örneklem, içinden alındığı evrendeki kadın oranı % 60, erkek oranı ise % 40 ise, yanlı olacaktır. Bu durumda kadınlar örneklemde fazla, erkekler ise eksik temsil edilmiş olacaktır. Örnekleme yanlılığının iki kaynağı vardır: Seçim yanlılığı ve tepki oranı yanlılığı. **Seçim yanlılığı**, örnekleme yapmak için kullanılan *işlemlerden* kaynaklanan ve evrenin bazı kesimlerinin fazla ya da eksik temsil edilmesiyle oluşur. Tepki oranı ile ilgili sorunları, "Anket Yöntemleri" bölümünde açıklayacağız.

Temsil ediciliği olan bir örneklemi oluşturan özellikler, incelenmek istenen evrene bağlıdır. Örneğin bir üniversite yönetimi okula arabalarıyla gelen öğrencilerin kampüsteki otopark alanı hakkında ne düşündüğünü öğrenmek istiyorsa, hedef evren kampüsteki tüm öğrenciler değil, okula arabalarıyla gelen öğrencilerdir. Bu durumda yanlı olmayan bir örneklem, kampüste arabaları olan öğrencilerin oluşturduğu evreni temsil eden bir örneklem olacaktır. Alıştırma I, seçim yanlılığından kaynaklanan yanlı örneklemleri belirlemeniz için hazırlanmıştır.

Örneklem Seçme Yaklaşımları

- Araştırma örneklemini seçmek için kullanılan iki yaklaşım, olasılığa dayalı örnekleme (probability sampling) ve olasılığa dayanmayan örneklemedir (nonprobability sampling).
- Olasılığa dayalı örnekleme, temsil edici bir örneklem oluşturmak için tercih edilen yöntemdir.

İLERİ ALIŞTIRMA I

Varsayımsal dört araştırma ve sonuçları ile aşağıdadır. Her bir örneklemin yanlı seçilip seçilmediğine karar verin ve yanlı seçildiğini düşünüyorsanız nedenini açıklayın.

1 Bir araştırmacı bölgesel psikoloji toplantısına kayıt yaptıran psikologların listesini alıyor ve bu listeden seçkisiz olarak oluşturduğu örneklemdeki herkesle tek soruluk bir görüşme yapıyor: "Psikoloji bilimi, zihinsel sorunları olan kişilerin belirlenmesine yardımcı olmak için üzerine düşeni yeterince yapıyor mu?" Araştırma bulgularına dayanarak "Psikologların çoğunluğu mental bozukluğu olan kişilerin belirlenmesine yardımcı olmak için psikolojinin üzerine düşeni yeterince yapmadığına inanıyor." şeklinde bir açıklama yapıyor.

2 Spor psikolojisi ile ilgilenen bir yüksek lisans öğrencisi, yerel bir spor takımının adında yapılacak değişikliğe karşı taraftarların tepkisini ölçmek istiyor. Bu amaçla taraftarlar sahadan ayrılırken seçtiği her 10'uncu kişiye yapılması düşünülen isim değişikliği ile ilgili görüşünü soruyor. Elde ettiği bulgulardan "Taraftarlar isim değişikliği istemiyor." sonucunu çıkarıyor.

3 Üniversite yönetimi, tüm birinci sınıf öğrencilerinin kampüste kalması planının uygulanıp uygulanmayacağına karar vermek için bir kampüs oylaması yapıyor. Heyecanlı bir okul gazetesi muhabiri, resmî olarak açıklanmadan önce sonuçlara ulaşmak istiyor. Birkaç arkadaşının da yardımıyla oy verme bölümünde bekleyerek çıkan öğrencilere ne yönde oy kullandıklarını soruyor. Oylama bittikten sonra aldığı cevapları sayıyor ve "kampüsteki öğrencilerin çoğunluğunun, birinci sınıftaki öğrencilerin kampüste yaşamalarını istediğini" bildiriyor.

4 Bir psikoloji profesörü, ergenlerde depresyon ve intihar konusuyla ilgili olarak kampüste düzenlenen film gösterimine katılacak öğrencilerine ekstra kredi vermeyi öneriyor. Gösterimin ardından bir asistan katılımcılara, bu şekilde ekstra kredi almaktan memnunlarsa ellerini kaldırmalarını söylüyor. On öğrenciden dokuzu ellerini kaldırıyor ve asistan profesöre, "Sınıfınızda hemen herkes bu şekilde ekstra kredi kazanmaktan memnun." diyor.

- Basit seçkisiz örneklemede, evrenin her birimi örnekleme girmek için eşit şansa sahiptir; tabakalı seçkisiz örneklemede evren alt evrenlere (tabaka) bölünür ve seçkisiz örneklem tabakadan çekilir.
- Olasılığa dayanmayan örnekleme (örneğin bulabildiğini örnekleme), evrendeki tüm birimlerin eşit seçilme şansını garantilemez.

Örnekleme yapmak için kullanılan iki yaklaşım vardır: Olasılığa dayalı örnekleme ve olasılığa dayanmayan örnekleme. Diyelim ki bir kolejdeki öğrencilerin kampüs güvenliği ile ilgili düşüncelerini öğrenmek istiyoruz. Kampüsteyken kendilerini ne kadar güvende hissediyorlar? Elbette bazı görüşlerin alınabilmesi için soruların ifade ediliş biçimi önemlidir. Ancak, şimdilik araştırmacının örnekleme yapmak için seçebileceği iki yaklaşımı düşünelim. İlkinde, kayıtlı tüm öğrencilerin yer aldığı listeden araştırmaya dahil etmek için 50 öğrenciyi seçkisiz olarak seçer. İkinci yaklaşımda, beş gün boyunca öğle yemeği saatinde okul kafeteryasına giderek araştırmasına katılacak 50 öğrenciye ulaşmaya çalışır.

Temel Kavramlar

İlk yaklaşımı **olasılığa dayalı örnekleme** olarak tanımlayabiliriz. Öğrencileri kayıt listesinden seçkisiz olarak seçerek listedeki her öğrenciye (birime), eşit oranda örnekleme girme şansı tanınmış olur. İkinci yaklaşım **olasılığa dayanmayan örneklemedir**. Katılımcıları, öğle yemeğini kafeteryada yiyen öğrenciler arasından seçmek, kayıt listesindeki herkesin eşit oranda örnekleme seçilme şansını garanti etmez. Listedeki her öğrencinin örnekleme girme şansı olmakla birlikte, her birimin seçilme olasılığının ne olduğunu tahmin edemeyiz.

KUTU 5.1

ANKET ARAŞTIRMASI İÇİN SEÇKİSİZ ÖRNEKLEM OLUŞTURMANIN AŞAMALARI

Adım 1: Araştırmak istediğiniz evreni belirleyin. Bütün mağaza çalışanları mı? Üniversitedeki part-time öğrenciler mi? Şehirdeki bir partiye kayıtlı olan kişiler mi? Bir üniversitedeki kimya bölümü mü?

Adım 2: Örnekleme havuzunuzu belirleyin. Örnekleminizi seçeceğiniz listeyi belirleyin; bir mağazada çalışanların tümünün listesi veya kayıt bürosundan temin edilecek part-time öğrencilerin listesi gibi. Örnekleme havuzunun her zaman ilgilendiğiniz evrene tam olarak karşılık gelmeyeceğini unutmayın. Örneğin, bir liste yakın zamanda güncellenmemiş olabilir ya da fark edilmeyen bazı hataları olabilir. Bu durumda ilgilendiğiniz evrenden olası sapmaların, araştırma sonucunu etkileyecek kadar ciddi olup almadığına karar vermeniz erekir.

Adım 3: Yönetilebilir büyüklükteki uygun bir listeye eriştiğinizde, temsil edici bir örneklem oluşturmak için seçkisiz örnekleme tekniği kullanılabilir. Bunu gerçekleştirmenin bir yolu seçkisiz numaralar tablosu kullanmaktır (bk. Ekler, Tablo A.1).

Pek çok durumda tam bir örnekleme havuzu oluşturulamaz. Çoğunlukla, bir evrendeki tüm potansiyel katılımcıların listesi yoktur veya potansiyel katılımcılara kolay ulaşılmaz veyahut liste etkin bir şekilde kullanılmayacak kadar uzundur. İnternet kullanıcıları? Bira içenler? Jogging yapanlar? Bekâr kadınlar? Gay çiftler? Silahı olanlar? Bu gibi evrenleri oluşturan katılımcıların listesine ulaşmanın zor olduğu durumlarda, örnekleme yöntemi olarak seçkisiz örnekleme en uygun sonucu verebilir veya araştırmacılar olasılığa dayanmayan örnekleme yöntemi kullanmak zorunda kalabilir. Bu seçenekler, sonraki bölümlerde ele alınmıştır. Göreceksiniz ki temsil edici bir örneklem elverişli olmasa bile önemli araştırmalar yapılmaktadır.

Pek çok öğrenci yemek için kafeteryayı tercih etmiyor olabilir, bazılarının o saatte dersi olabilir ya da şehirlerarası gidip gelen bazıları yemeklerini başka bir yerde yiyor olabilir.

Seçilen örneklemin evreni temsil etmesini garantilemek için olasılığa dayalı örnekleme, olasılığa dayanmayan örneklemeden çok daha güçlü bir yaklaşımdır. Kayıtlı öğrenci listesinden seçkisiz olarak seçilen 50 öğrencinin, kampüsteki öğrencileri temsil etme olasılığı kafeteryada yemek yerken seçilenlerden daha yüksektir.

Temel Kavram

Olasılığa Dayalı Örnekleme Olasılığa dayalı örneklemenin belirgin özelliği, araştırmacının, evrendeki her birimin örnekleme seçilme olasılığını belirleyebilmesidir. Olasılığa dayalı yapmak için kullanılan iki yaygın teknik, basit seçkisiz örnekleme ve tabakalı seçkisiz örneklemedir. Basit seçkisiz örnekleme, olasılığa dayalı örneklemenin temel tekniğidir. **Basit seçkisiz örneklemenin** en yaygın tanımı, bütün birimlerin örnekleme girme olasılıklarının eşit olmasıdır. Basit seçkisiz örnekleme yapmak için kullanılan işlemler Kutu 5.1' de özetlenmiştir.

Seçkisiz bir örneklem oluştururken karar verilmesi gereken önemli bir husus, örneklemin büyüklüğüdür. Şimdilik şunu bilmek yeterli olacaktır: Evreni temsil edecek seçkisiz örneklemin büyüklüğü, evrendeki çeşitliliğe bağlıdır. Örneğin, Ivy League okullarında okuyan öğrenciler, akademik yetenekleri bakımından, Amerika'daki tüm kolej öğrencilerinden daha homojen bir evreni temsil eder. Bir uçta en homojen evren, tüm üyeleri ilgili tüm özellikler bakımından özdeş olan evrendir. Bu durumda tek birimden oluşan bir örneklem bile evrenin büyüklüğü ne olursa olsun bu evreni temsil eder. Diğer uçta ise en heterojen evren, her üyesi birbirinden tamamen farklı olan bir evren olacaktır. Bu durumda,

İLERİ ALIŞTIRMA II

İki öğrenci araştırmacı, kampüsteki kız öğrenci topluluklarına ve erkek öğrenci topluluklarına yönelik tutumları belirlemek için bir anket uyguluyor. Okulda 3200 öğrenci var. Öğrencilerin %25'i Yunan organizasyonuna bağlı, %75'i ise değil. Bu iki öğrenci araştırmacı bu çalışma için en iyi örnekleme planı üzerinde anlaşamıyorlar. Araştırmacılardan birisi 100 katılımcıyı Yunan organizasyonundan, 100 tanesini ise bağımsız olan öğrenciler arasından olacak şekilde, tabakalı seçkisiz örnekleme yöntemi kullanmak istiyor (toplam 200 kişi). Diğeri ise kampüsün tamamından basit seçkisiz örnekleme yöntemi ile 100 kişi almak istiyor.

1 Bu iki örnekleme planının ne ölçüde temsil edici olduğunu ve Yunan organizasyonuna dahil öğrencilerin görüşlerinin güvenilir bir şekilde ölçülmesi ihtimalini eleştirel olarak yorumlayın.

2 Eğer ikisinin de uygun olmadığına karar verirseniz kendi örnekleme planınızı geliştirin.

büyüklüğü ne olursa olsun hiçbir örneklem bu evreni temsil edemez. Böyle heterojen bir evreni temsil etmek için, evrenin her bireyini örnekleme dahil etmek gerekir. Uygulamada, araştırmacılar bu iki uç durumun ortasında çalışır.

Temel Kavram

Bir örneklemin temsil ediciliği, tabakalı seçkisiz örnekleme yöntemi kullanılarak geliştirilebilir. **Tabakalı seçkisiz örnekleme** yönteminde evren, tabakalar adı verilen alt evrenlere bölünür ve seçkisiz örneklemler bu tabakalardan çekilir. Her tabakadan ne kadar birimin seçileceğini belirlemek için genellikle iki yöntem kullanılır. İlk yöntemde her tabakadan eşit büyüklükte örneklem çekilir. İkinci yöntemde ise birimler oranlarına göre seçilir. Örneğin, % 30'u birinci, % 30'u ikinci, % 20'si üçüncü ve % 20'si son sınıf öğrencilerinden oluşan bir evren düşünelim (bu örnekteki tabakalar, sınıflardır). Bu evrenden çekilecek 200 kişilik tabakalı seçkisiz örneklemde, 60 birinci, 60 ikinci, 40 üçüncü ve 40 dördüncü sınıf öğrencisi yer alacaktır. Her tabakadan eşit sayıda çekilerek oluşturulacak örneklemde ise her sınıftan 50 öğrenci olacaktır. *Bu durumda sadece oranlı tabakalı örneklem temsil edicidir.*

Tabakalı seçkisiz örnekleme, örneklemlerin temsil ediciliğini artırma potansiyelinin yanı sıra evrenin özel oranlarını belirlemek istediğinizde de yararlıdır. Örneğin, 2000 öğrencili bir kampüsteki öğrenci tutumlarını ölçmek için yapılacak bir araştırmada, 100 öğrenciden oluşan basit seçkisiz örneklem yeterli olacaktır. Bununla birlikte, varsayalım ki örnekleminizde kampüsteki 40 kimya öğrencisinden sadece 2'si yer alıyor ve siz de farklı bölümlere göre öğrencilerin görüşlerini belirlemek istiyorsunuz. Bu dağılım, kampüsteki kimya bölümünü doğru oranda yansıtsa bile, 40 kimya bölümü öğrencisini temsilen sadece iki öğrencinin görüşünü almak riskli olur (2 çok düşük bir sayıdır). Bu durumda, (ve genellikle bir tabaka sayıca küçük olduğunda) kimya bölümünün görüşlerini daha iyi belirlemek için örnekleme daha fazla kimya bölümü öğrencisi alabilirsiniz. Daha önce de belirttiğimiz gibi, örneklemimize tam olarak kaç kişiyi alacağımızı söyleyemeyiz, çünkü temsil edici örneklemin büyüklüğü, evrenin çeşitliliğine bağlıdır.

Olasılığa Dayanmayan Örnekleme Olasılığa dayanmayan örneklemenin en bilinen türü, *bulabildiğini örneklemedir*. Bulabildiğini örneklemede katılımcılar öncelikle erişilebilirliklerine ve gönüllü katılım isteklerine göre seçilir. Örneğin, gazeteler genellikle "sokaktaki insan"ın görüşlerini yayımlarlar. Bunları okumak ilginç olabilir ancak bu insanların görüşleri daha büyük bir topluluğun görüşlerini temsil etmeyebilir. Temsil edicilikle ilgili zafiyet oluşur, çünkü bulabildiğini örnekleme, olasılığa dayanmayan örneklemedir ve biz, toplumdaki her bireyin bu örnekleme girme şansı olup olmadığından emin olamayız. Dergiler aracılığıyla yürütülen araştırmalardaki örneklemler de bulabildiğini örneklemedir. Çünkü hem dergiler erişilebilir olmalı (satın alınabilir) hem de katılımcılar cevaplarını gönüllü olarak göndermelidir. Üniversitelerde birçok psikoloğun kullandığı ve genellikle psikolojiye giriş dersini alan kayıtlı öğrencilerden oluşan "katılımcı havuzu", bulabildiğini örnekleme tekniğine bir örnektir.

Crossen (1994), başka bir bulabildiğini örnekleme tekniği olan telefon anketlerinin sakıncalarına değiniyor. Telefon anketleri, televizyon ve radyo programları tarafından hedef kitlelerinin görüşlerini almak için kullanılır. Programı takip edenler ve telefon etmeye gönüllü olanlar (900'lü numarayı aramak için bazen ücret ödemek gerekir) bu telefon anketlerinin örneklemini oluşturur. Ankete katılmak için telefon edenler evrenin genelinden, sadece belli bir programı izledikleri için değil, aynı zamanda telefon edecek kadar motive oldukları için de farklıdırlar. Benzer şekilde, ekranda beliren bir anket sorusuna cevap veren bilgisayar kullanıcıları, cevap vermemeyi tercih eden kullanıcılardan (ya da bilgisayar kullanıcısı olmayanlardan) farklıdırlar.

Bir prime time TV haber programı, düzenlediği telefon anketi ile izleyicilerine "Birleşmiş Milletlerin merkezi Amerika Birleşik Devletleri'nde kalmalı mıdır?" sorusunu sormuştur (Crossen, 1994). Sonradan ortaya çıkmıştır ki başka bir anket araştırmasında da aynı soru, seçkisiz olarak belirlenen 500 kişiye sorulmuştur. Telefon anketine cevap veren 186,000 kişinin çoğunluğu (%67), BM'nin Amerika'dan taşınması yönünde görüş bildirmiştir. Öte yandan, diğer anket araştırmasına cevap veren 500 kişilik katılımcı grubunun büyük çoğunluğu ise (%72) BM'nin Amerika'da kalmasını istemiştir. Nasıl olur da bu iki araştırmadan birbirinden tamamen farklı sonuçlar çıkabilir? Daha büyük örneklemi olduğu için telefon anketinin sonuçlarına daha fazla mı güvenmeliyiz? Kesinlikle Hayır! Büyük bir örneklem de herhangi bir kadar temsil edicilikten uzak olabilir. Genel kural olarak şunu hatırda tutmakta yarar var: *Bulabildiğini örnekleme tekniğini kullandığınızda, örneklemin temsil ediciliğini gösteren güçlü kanıtlar yoksa, örnekleminiz yanlı olacaktır.*

ANKET YÖNTEMLERİ

- Anketle veri elde etmek için dört yöntem kullanılır; posta anketleri, yüze yüze görüşmeler, telefon görüşmeleri ve internet anketleri.

Örneklem seçmek, anketle araştırma yaparken verilmesi gereken çeşitli kararlardan yalnızca birisidir. Bazı durumlarda, katılımcılara kolaylıkla ulaşılabildiği için anket yöntemiyle araştırma yapmak kolaydır. Bir sınıftaki öğrencilerden anketi cevaplamaları veya dinî bir topluluğun üyelerinden ibadet

hizmetleri ile ilgili bir listedeki maddeleri cevaplamaları istenebilir. Bu gibi duruma özel anketlerde çoğu kez anketin yapıldığı anda bulunan kimselerden oluşan örneklem ile kâğıt-kalem ölçümleri kullanılır. Diğer anket yöntemleri ise belirli bir durum ya da ortamla sınırlı değildir ve potansiyel bir katılımcı evreninden örneklem seçme işlemleri kullanabilir. Ayrıca, katılımcılardan ne şekilde cevap alacağınıza da karar vermeniz gerekir. Anket yoluyla veri elde etmek için dört yöntem kullanılır; yüz yüze görüşmeler, telefon görüşmeleri, internet anketleri ve posta anketleri. Araştırma yaparken her zaman doğru olan ve bütün durumlar için en iyi olan bir anket yöntemi yoktur. Her yöntemin kendi üstünlükleri ve dezavantajları vardır. Araştırma yaparken zor olan, kendi araştırma sorunuza en uygun yöntemi seçmektir.

Posta Anketleri

- Posta anketleri hızlı ve pratik olmakla birlikte, katılımcıların anketleri doldurmadıkları ve geri dönmedikleri zaman geri dönüş oranları düşmektedir.
- Geri dönüş oranı ile ilgili sorunlar nedeniyle, araştırmaya katılan örneklem söz konusu evreni temsil etmeyebilir.

Posta anketleri, kendi kendine uygulanan yani katılımcıların soruları kendilerinin yanıtladıkları anket formlarıdır. Posta anketlerinin bir üstünlüğü, görece olarak hızlı tamamlanmalarıdır. Kendi kendine uygulanabildiğinden görüşmeci yanlılığından (sonraki bölümde açıklanacaktır) kaynaklanan problemleri de engeller. Bu dört anket yöntemi arasında, posta anketleri, özellikle katılımcılar gerçek isimlerini kullanmadıkları durumlarda, kişisel ve katılımcıyı utandırabilecek konuların araştırılabilmesi için en iyi yöntemdir.

Ne yazık ki posta anketlerinin bir çok dezavantajı da vardır. Bir zorluğu, bütün katılımcılar için bir kopya hazırlama ve postalama maliyetidir. Aynı zamanda katılımcılar anketle ilgili soru soramayacakları için kullanılan sorular tüm katılımcıların kolayca anlayabilecekleri şekilde çok açık olmak zorundadır. Daha az önemli bir dezavantaj ise araştırmacıların katılımcıların soruları cevaplama sırası üzerinde kontrolünün olmamasıdır. Soruların sırası katılımcıların belli sorulara nasıl cevap verdiklerini etkileyebilir. Ama bunların arasında en ciddi problem, posta anketlerinde yanlılığa neden olabilecek düşük geri dönüş (tepki) oranıdır.

Temel Kavram

Tepki oranı, anketi tamamlayanların yüzdesini ifade eder. Örneğin, 100 kişiden 30'u anketi tamamlamışsa tepki oranı %30'dur. *Düşük tepki oranı, örneklemin temsil ediciliğini tehdit eden bir* **tepki oranı yanlılığı** *olabileceğine işaret eder.* Bu duruma neden olan birçok etken olabilir. Örneğin, okuma-yazma sorunu olan veya eğitim düzeyi düşük olan veyahut görme sorunları olan katılımcılar anketi tamamlayamayabilirler. Bu nedenle, bu gibi özellikleri olan katılımcılar son örneklemde iyi temsil edilemeyebilirler. Genellikle, rastgele seçilen insanlar ya çok meşgul olabilir ya da araştırma sorularını tamamlayıp göndermekle yeterince ilgilenmeyebilirler. Düşük tepki oranı örneklemin, araştırılmakta olan evreni temsil etmemesine neden olabilecek en önemli etmendir. Bu yüzden, dikkatlice belirlenmiş olasılığa dayalı bir örnekleme, olasılığa dayanmayan bir örneklemeye –bulabildiğini örneklemeye– dönüşebilir; yani, sadece

ulaşılabilen ve anketi doldurmaya istekli olan katılımcıların görüşlerini temsil eden bir örnekleme dönüşebilir.

Araştırmanın en başında örneklem çok özenle seçilse bile ve tepki oranı %100 olsa bile araştırma, potansiyel olarak tepki oranı yanlılığı tehlikesi taşıyabilir. Bununla birlikte, düşük tepki oranı otomatik olarak örneklemin evreni temsil etmediğini göstermez. Araştırmacı, anketi cevaplayan katılımcılardan oluşan son örneklemin ne ölçüde evreni temsil ettiğini, evrenin fazla veya eksik temsil edilen bir bölümünün olup olmadığını göstermek zorundadır.

Posta anketlerinde tipik olan tepki oranı %30 civarındadır. Ancak bu oranı artırmak için yapılacak bazı şeyler vardır:

- Soru listesinin kişisel bir havasının olması (örneğin, yalnızca "yurtta kalanlar"a veya "öğrenci"ye değil, katılımcının kendisine hitap etmesi);
- Soruların kolay cevaplanabilir olması;
- Anket konusunun katılımcıya ilginç gelmesi;
- Katılımcıların, araştırmayı yapan ve destekleyenlerin kim olduğunu anlayabilmesi.

Yüz Yüze Görüşmeler

- Maliyeti yüksek olmakla birlikte, yüz yüze görüşmelerde araştırmacılar, anketin uygulanması üzerinde daha fazla kontrol kazanır.
- Görüşmeciler yanıtları yanlış kaydettiği zaman veya katılımcının cevaplarını yönlendirdiği zaman görüşmeci yanlılığı ortaya çıkar.

Anket verisi toplamak için yüz yüze görüşmeler kullanıldığında, katılımcılarla genellikle evlerinde, alışveriş merkezlerinde iletişim kurulur ve form daha önceden eğitim almış anketörler tarafından uygulanır. Yüz yüze görüşmeler posta anketlerine kıyasla soru sorarken esneklik sağlar. Görüşme sırasında katılımcılar, soruları anlamadıkları zaman açıklama isteyebilir; ayrıca, önceden eğitim almış olan görüşmeciler, açık uçlu sorulara verilen cevaplar belirsiz olduğu zaman, katılımcıdan ek açıklama isteyebilir. Görüşmeci, soruların sırasını takip eder ve katılımcıların tüm katılımcıların soruları aynı sıra ile cevapladığından emin olabilir. Tipik olarak, yüz yüze görüşmelerde katılım oranı posta anketlerine göre daha yüksektir.

Yüz yüze görüşmelerin kullanılmasının avantajları dikkat çekicidir, ama birkaç dezavantajı da vardır. Kentteki suç oranı ve gün içerisinde sokakta olanların sayısı arttığı (evde duranların sayısı azaldığı) için, evlerde yapılan yüz yüze görüşmelere olan ilgi azalmıştır. Yüz yüze görüşme yapmanın en belirgin dezavantajı, maliyetli olmasıdır. Eğitimli görüşmeci kullanılması hem zaman hem de para açısında pahalıya mal olur. Belki en önemli dezavantaj, yüz yüze görüşmelerde görüşmeci yanlılığının olası olmasıdır. Görüşmeci soruları ve cevapları aktarırken tarafsız kalabilmelidir. **Görüşmeci yanlılığı**, anketör katılımcıların yanıtlarının sadece bir kısmını kaydettiğinde veya "uygun" olacak şekilde cevapları değiştirdiğinde ortaya çıkar. Örneğin, televizyon kanalları ile ilgili bir araştırmadaki bir katılımcıyı düşünelim. "Televizyon programları ile ilgili en büyük problem çok fazla şiddet içermesidir." Eğer görüşmeci katılımcının bu cevabının tamamını yazmak yerine sadece " TV şiddeti" olarak yazarsa görüşmeci yanlılığı yapmış olur. Bir sonraki "Şiddet derken cinayet ve tecavüzü mü

Temel Kavram

kastettiniz? " diye sorduğunda da yine görüşmeci yanlılığı yapılmış olunur. Daha tarafsız soruşturma, katılımcının ne söylemek istediğini kendisine sormaktır: "Şiddet derken neyi kastettiğinizi açıklar mısınız?" gibi.

Görüşmeci yanlılığına karşı en iyi önlem, soruları kelimesi kelimesine takip edecek, yanıtları aynı şekilde kaydedecek ve izleme sorularını sağduyulu kullanacak şekilde eğitilmiş görüşmecilerin iyi bir maaşla motive edilerek çalıştırılmasıdır. Aynı zamanda görüşmecilere, karşılaştıkları sorunlarla ve kafa karıştırıcı durumlarla nasıl başa çıkacaklarını detaylı olarak açıklayan yönergeler verilmelidir. Son olarak, görüşmeciler araştırma projesinin yöneticisi tarafından sıkı bir şekilde denetlenmelidir.

Bilgisayar teknolojisi yüz yüze görüşme ve kendi kendine uygulanan anketlerin bir melezini kullanmayı mümkün kılmıştır. Bir katılımcı, görüşmecinin bilgisayarda okuduğu soruları yine bilgisayarda cevaplayabilmektedir. Bu teknoloji ile birlikte her katılımcı soruları tam anlamıyla aynı görüşmeci tarafından aynı şekilde okunmuş hâliyle duymaktadır. Bu da görüşmeci yanlılığı riskini düşürmektedir. Bu teknoloji aynı zamanda çok kişisel olan soruların daha gizli bir şekilde cevaplanmasına izin vermektedir (Rasinski, Willis, Baldwin, Yeh & Lee, 1999).

Telefon Görüşmeleri

- Bazı dezavantajları olmasına rağmen telefon görüşmeleri kısa anketler için sıklıkla kullanılmaktadır.

Kişisel görüşmelerin masraflı olması ve görüşmecilerin denetlenmesindeki güçlükler anket araştırmacılarının telefon ve internet anketlerine dönmelerine neden olmuştur. Telefon numaralarını rastgele çeviren teknikler, araştırmacıların telefon kullanıcılarından oluşan temsil edici örneklem ile iletişime geçmelerine olanak sağlamıştır. Telefon görüşmeleri tehlikeli mahallelere, kilitli binalara ve sadece akşamları uygun olan katılımcılara daha iyi erişim sağlamıştır (daha önce hiç akşam yemeği sırasında bir telefon anketi uyguladınız mı?) Görüşmeler telefonla daha hızlı tamamlanmaktadır, ve bütün görüşmeler bir yerden yapıldığı için görüşmeciler daha iyi denetlenebilmektedir (Şekil 5.2).

Diğer anket yöntemlerinde olduğu gibi, telefon anketleri de kusursuz değildir. Telefonu olanların katılımı nedeniyle olası seçme yanlılıkları ve görüşmeci yanlılıkları devam etmektedir. Katılımcıların hangi uzunlukta telefonda kalmak isteyecekleri bir sınırlılıktır ve "kimliği belirsiz" bir kişi tarafından sorulan sorulara verilen cevaplar, kişisel görüşmelerde verilecek cevaplardan farklı olabilir. Birçok insanın sabit hatlar yerine sadece cep telefonu kullanıyor olması, insanların nerede bulundukları ve ne zaman anketi yanıtladıkları gibi ortam değişkenlerinden kaynaklanan potansiyel bir tepki farklılığı getirecektir. Ayrıca, yüksek sosyo-ekonomik seviyedeki insanların birden fazla telefon numarası kullanmaları, numaraların rastgele çevrildiği teknikle yapılan araştırmalarda, aynı kişinin birden fazla sayıda katılmasına neden olabilir. Hippler ve Schwraz (1987) insanların telefon anketlerinde karar vermeye az zaman ayırdığını ve görüşmecinin söylediği cevap seçeneklerini hatırlamakta zorluk çekebileceklerini belirtmiştir. Ayrıca, ürün satmak ve bağış toplamak için telefon aramalarının fazla yapılması, birçok insanda görüşme yapma konusunda isteksizliğe yol açmıştır.

ŞEKİL 5.2 Kısa anketler için telefon numaralarını rastgele çevirme yöntemi, genel olarak telefon kullananları temsil eden örnekleme etkili bir şekilde erişim sağlar.



Görüntülü konuşma ve sesli mesaj gibi seçenekler insanların istenmeyen çağrılardan kaçınmalarını kolaylaştırmıştır. Bütün bu sınırlılıklarına rağmen telefon görüşmeleri belki de sanıldığından daha çok, kısa anketler için sıklıkla kullanılmaktadır.

İnternet Anketleri

- İnternet, yaygın, potansiyel olarak çeşitli ve yeterince temsil edilmemiş örneklemlerden gelen anket verisi sağlayan etkili ve düşük maliyetli bir yöntem olduğu için anket araştırmalarına çok sayıda avantaj sağlamıştır.
- İnternet anketleri kullanan araştırmaların dezavantajı, potansiyel tepki oranı yanlılığı ve seçme yanlılığı içermesi ve araştırma ortamının kontrol edilememesidir.

Anketler, internete dayalı davranışsal araştırmaların en eskileri arasındadır. Katılımcılar, bir soru listesini online olarak doldururlar ve "gönder" tuşuna bastıklarında tepkileri kaydedilmiş olur. Yazılımlar giderek daha karmaşık hâle geldiği için, milyonlarca tepki otomatik olarak kaydedilebilmekte ve işlenebilmektedir. Programlar aynı zamanda değişkenlerin değişimlenmesine ve katılımcıların deneysel koşullara seçkisiz atanmalarına olanak sağlamaktadır [Örnek olarak Fraley, 2004 ve Gosling ve Johnson, 2010; bir rehber olarak internette HTML-temelli psikolojik araştırmalar'a bakınız ve yararlı internet kaynakları için Birnbaum, 2004 ile Kraut ve arkadaşlarına (2004) bakınız].

İnternet araştırmalarının sayısız üstünlüğü hemen akla gelecektir. Listenin en üstünde, etkili ve düşük maliyetli olması vardır [örneğin, Buchanan, 2000; Skitka ve Sargis'e (2005) bakınız]. Milyonlarca değilse bile binlerce farklı yaş, etnik köken ve ulusa mensup katılımcıyla bilgisayarda bir kaç tuşa dokunarak bağlantı kurabiliriz. Zaman ve iş gücü sarfı, posta veya telefon anketleriyle ve elbette kişisel görüşmelerle kıyaslandığında çarpıcı şekilde düşüktür. İnternetteki soru listeleri kağıda ihtiyaç duymaz ve böylece kopyalama masrafları ve doğal kaynak kullanımından tasarruf edilir. Katılımcılar uygun oldukları zaman ve evlerinin, işyerlerinin, yurtlarının sunduğu rahatlık içerisinde cevaplayabilirler. Ayrıca, araştırmacılar internet anketlerine kolaylıkla video, görsel ve işitsel ipuçları ekleyebilir ve cevap seçeneklerini katılımcının sorulara cevap vermesine göre özelleştirebilir. Soruların sırası ve cevap seçenekleri, tepki yanlılığını azaltmaya yardımcı olacak şekilde seçkisiz olabilir (Tuten, 2010).

Geniş ve potansiyel olarak çeşitli bir örnek grup üzerinde çalışmaya ek olarak, internet anketleri psikolojik araştırmalarda tipik olarak düşük temsil edilen gruplara ulaşmayı da sağlayabilir (Skita ve Sargis, 2005). Web ortamındaki sohbet odalarının yaygınlığı, özel-ilgi grupları ve destek gruplarının içine girebilmek, araştırmacıya şu gibi özel gruplara ulaşabilme olanağı sağlar: Evcil hayvan sahipleri, nefret grubu üyeleri, kanserle mücadele edenler, suç kurbanları veya geleneksel anket yöntemleri ile kolayca araştırılması mümkün olmayan bazı özel katılımcı grupları. İnternet gerçekten tam olarak dünya çapında bir katılımcı kaynağı olduğu için, kültürler arası araştırmalar için de yeni olanaklara kapı açmaktadır (örneğin, Gosling ve ark., 2004).

Kuşkusuz, internette yapılan araştırmaların da dezavantajları vardır. Bu dezavantajların en başında, potansiyel bir yanlı örneklem tehlikesi gelir (Birnbaum, 2000; Kraut ve ark., 2004; Schmidt, 1997). Bugün için, internet kullanıcılarından seçkisiz bir örnek grup elde etmenin bir yöntemi yoktur. Sonuç olarak internet araştırmacıları, olasılığa dayanmayan örneklemlerle (nonprobability sample) yetinmek zorundadır. Hem tepki oranı yanlılığı ve hem de seçim yanlılığı olması muhtemeldir. Düşük tepki oranından kaynaklanan sorunlar, diğer anket yöntemlerindeki gibi cevap vermeme nedeniyle oluşabilir. Aslında internet anketlerindeki tepki oranı, tipik şekilde, posta ve telefon anketlerindeki tepki oranından daha düşüktür [Kraut ve ark., 2004; Skita ve Sargis'e (2005) bakınız]. Daha önce gördüğümüz gibi, bir ankete cevap veren bireyler, cevap vermeyenlerden önemli bazı özellikleri bakımından farklıdır. Seçme yanlılığı, örneklem grubun internete girebilen bireyler arasından seçilmesinden kaynaklanır. Birleşik Devletler'deki yüksek gelirli hane halkının interneti daha çok kullanması muhtemeldir veya çocuklu haneler, çocuksuz hanelerden daha çok internet kullanıyor olabilir. Beyaz ve Asyalı aileler, siyah veya İspanyol kökenli ailelerden iki kat daha fazla internete girmektedir (Newburger, 2001).

Seçme yanlılığı, katılımcı peşinden koşma nedeniyle daha da artabilir. Araştırmacılar, araştırmalara katılımı arttıran (örneğin, Bölüm 1'de sözü edilen APS'nin web sitesi gibi) internet sitelerinde araştırma ilanları yayımlayarak örneklem grup oluşturabilirler veya kendi anketleri (örneğin, *kişilik anketi*) için bir web sayfası kurarak kullanıcıların o sayfayı ziyaret etmelerini bekleyebilirler (Krantz ve Dalal, 2000). Günümüzde şirketler, katılımcı "panellerine" üye yaptıkları potansiyel katılımcılara ücret karşılığında erişim olanağı

KUTU 5.2

İNTERNETTE ANKET YAPMA

İnternet anket araştırmaları için yaygın bir katılımcı kaynağıdır. Hem üstünlükleri ve hem de dezavantajları vardır. Özellikle belirtmeliyiz ki her kim internette araştırma yürütmekle ilgileniyorsa, önce bu yöntemle bağlantılı etik ve işlemsel meseleleri öğrenmelidir. Metin içinde yararlı inceleme makaleleri ve bazı kaynaklar verilmiştir. Ve araştırmaya başlamadan önce, *internet anket araştırmalarında deneyimli bir araştırmacının önerilerini alın*.

İnternet anket araştırmacılarının başlangıçta karşılaştığı iki büyük engel, uygun bir örneklemin seçilmesi ve anketin uygulanmasıdır.

- **Örneklem Seçme** İnternet kullanıcıları arasından seçkisiz bir örneklem seçme mümkün olmamakla birlikte, tipik üniversite öğrencilerinden çok daha çeşitli olan bir "katılımcı havuzu" oluşturmak için bazı seçenekler mevcuttur. Araştırmanızın amacına erişmek için bir ya da daha fazla örneklem seçmek makul olabilir, ama bu seçeneğin pahalı olabileceğini de belirtmeliyiz.
- **Anketin Uygulanması** "Uygulama" ile kast ettiğimiz şey, bir anketin internette kullanılabilecek şekilde geliştirilmesi, kaydedilmesi ve cevapların düzenlenmesidir. Anketin başladığı noktadan itibaren ayrıntılı ve yeterli bir veri toplama aracı hâline gelmesi için olduğu kadar, veri toplama sürecinin etkili olabilmesi için de bazı programlama bilgileri gereklidir. Örneğin, multimedya uygulamaları, soruların ve cevap seçeneklerinin yansız dizilmesi, katılımcıya geri bildirim verilmesi için kullanılan programlar. Yine de basit bir internet anketi yürütmek, interneti çok az bilen veya hiç bilmeyenler için bile eskisi kadar göz korkutucu bir iş değildir. Çeşitli şirketler (ücretli olarak) kullanımı kolay yazılımlar ve destek hizmeti sağlamaktadır

sağlamaktadır [örneğin, Tuten'e (2010) bakınız) veya web siteleri çeşitli denek havuzlarında yer alan katılımcılara ücret olanağı sağlıyor (örneğin, Amazon's Mechanical Turk; ayrıca Buhrmester, Kwang ve Gosling, 2011'e bakınız). Diğer stratejiler, araştırma konusuna ilgi duydukları için katılmak isteyen bireylere veya gruplara araştırma ilanlarını göndermeye yöneliktir. Bununla birlikte, Skitka ve Sargis'in (2005) vurguladığı gibi, nasıl ki internet kullanıcıları genel evreni temsil etmiyorsa gibi, aynı şekilde, özel internet ilgi grupları da kendi özel gruplarını temsil etmeyebilir.

Araştırma ortamını kontrol eksikliği de internet anketlerinin temel dezavantajlarından birisidir (Birnbaum, 2000; Kraut ve ark., 2004). Bölüm 3'te işaret ettiğimiz gibi kontrol eksikliği, bireyi araştırmaya katıldığı için uğrayabileceği muhtemel zararlardan koruma (örneğin, anket sorularından kalan duygusal üzüntüler) ve bilgilendirilmiş onamla ilgili ciddi etik sorunları arttırır. Araştırmacı ortamda mevcut olmadığı için, katılımcının yönergeyi iyi anlayıp anlamadığını, cevaplarını ciddiyetle, sağduyulu ve iyi niyetle verip vermediğini veya birden fazla kayıt yapıp yapmadığını tespit etmenin kolay bir yolu yoktur (Kraut ve ark., 2004). Katılımcılar, dikkati dağıtan bir ortamdayken, tek başlarına veya grup hâlindeyken anketi cevaplayabilirler (Skita ve Sargis, 2005). Bir internet araştırmacısı, olasılık ve risk konulu bir anketi cevaplamak için katılımcıların, yönergede kullanmamaları belirtildiği hâlde, hesap makinesi kullanmalarından endişeye kapılmıştır (Birnbaum, 2000). İnternet araştırmalarının üstünlüklerinin çok sayıdaki dezavantajına ağır basacağını söylemek mümkün görünüyor. Teknolojinin geliştikçe ve etik kurullar insan katılımcıları koruyan, kabul edilebilir yöntemler tasarladıkça internet anketleri, bir anketle veri toplama yöntemi olarak gelişmesini sürdürecektir (Kutu 5.2'ye bakınız).

ANKET ARAŞTIRMA DESENLERİ

- Anket desenlerinin kesitsel desen, boylamsal desen ve ardışık bağımsız örnekleme deseni olmak üzere üç türü vardır.

Bir anket araştırmacısının vermek zorunda olduğu en önemli karar, araştırma deseninin seçilmesidir. Bir anket araştırması deseni, tüm çalışmanın yürütülmesi için kullanılan plandır. Üç tür anket araştırma deseni vardır: Kesitsel (cross-sectional design) desen, ardışık bağımsız örneklemler deseni (successive independent samples design,) ve boylamsal (longitudinal design) desen. Her amaca uygun olan bir anket araştırması deseni yoktur. Araştırmacı, çalışmasının amacına bağlı olarak bir desen seçer.

Kesitsel Desen

- Kesitsel desende, evrenden eşzamanlı olarak bir veya daha fazla örneklem alınır.
- Kesitsel desen, evrenin özelliklerini veya iki veya daha fazla evren arasındaki farkları betimlemeye olanak sağlar ve kesitsel desenle elde edilen korelasyonel bulgular, araştırmacının tahmin yapmasını sağlar.

Temel Kavram

Kesitsel desen en yaygın kullanılan anket araştırması desenlerinden biridir. **Kesitsel bir desen**de, evrenden bir veya daha fazla örneklem *eş zamanlı* olarak alınır. Bu desenin odağı betimlemedir –belirli bir zamanda iki veya daha fazla evren arasındaki farkları veya bir evrenin özelliklerini betimleme. Örneğin, kesitsel bir desende kız ve erkek üniversite öğrencilerinin, kendilerini Facebook sayfalarında "ilişkisi var" olarak bildirmelerine yönelik tutumları araştırılmıştır (Fox ve Warber, 2013). Araştırmacılar bir evrendeki bir örneklemi (bir "kesit") belirli bir zaman diliminde araştırdıkları için desen kesitseldir. Ek ders kredisi karşılığında internetteki ankete katılmak üzere, Batı'daki büyük bir üniversiteden 403 katılımcı toplamışlardır. Araştırmacılar, öğrencilerden "Facebook'ta ilişkisi var görünüyorsa, iki partnerin birbirleriyle çıktığını düşünürüm" veya "Bir kişi Facebook'ta ilişkisi var görünse bile, yine de başkalarıyla çıkabileceğini düşünürüm" gibi ifadeler için 5'li ölçek (1: Kesinlikle katılmıyorum, 5: Kesinlikle katılıyorum) kullandılar (s. 5). Bu anketin bulguları, Facebook'ta "ilişkisi var" olarak belirtilmiş olan heteroseksüel ilişkilerde dikkat çekici hususları işaret etti. Örneğin, Facebook'ta ilişkisi var görünmenin "ayrıcalık ve ciddiyet" anlamına geldiğine, erkekler kadınlara göre daha az inanmaktadır (s. 6).

Olasılığa dayanmayan örneklem kullanan anket araştırmacıları, örnekleme yöntemlerinin, örneklemin temsil ediciliğini ve bulguların genellenebilirliğini nasıl etkilediğini açıklamalıdır. Facebook'taki romantik ilişkileri araştıran araştırmacılar, örneklemlerinin tüm Facebook kullanıcılarını temsil etmediğine, örneğin gay ve lezbiyen kitleleri ve Amerika dışındaki öğrencileri yeterince temsil etmediğine dikkat çekmişlerdir (Fox ve Warber, 2013). Aynı zamanda bu araştırmacılar, çalışmalarının "keşif amaçlı" olduğunu da belirtmişlerdir. Bir araştırmacı veya okuyucu olarak değerlendirilmesi gereken en önemli şey, araştırma örnekleminin, araştırılan evreni ne derecede temsil ettiğidir. Bunu yapmanın en

iyi yolu örnek seçme yönteminin incelenmesidir: Örneklem, araştırılan evrenle büyük ölçüde eşleşen, iyi tanımlanmış bir örnekleme havuzundan seçkisiz olarak seçildi mi? Eğer bu şekilde seçilmediyse, araştırma bulgularından hangi sonuçlara varılabileceği konusunda titiz olmayı bilmek gerekir.

Kesitsel yöntem, anket araştırmalarının betimsel ve tahmine yönelik amaçlarına en uygun araştırma desenidir. Anketler ayrıca, zaman içindeki tutum ve davranış değişikliklerini ölçmek ve doğal şekilde ortaya çıkan olayların etkilerini belirlemek için de kullanılır. Bu amaçlar için seçilecek yöntem, kesitsel desen değildir. Tercihen, denekleri zaman içinde sistematik şekilde örneklem olarak alan araştırma desenlerine ihtiyaç vardır. Bu tür iki desen sonraki iki bölümde ele alınmıştır.

Ardışık Bağımsız Örneklemler Deseni

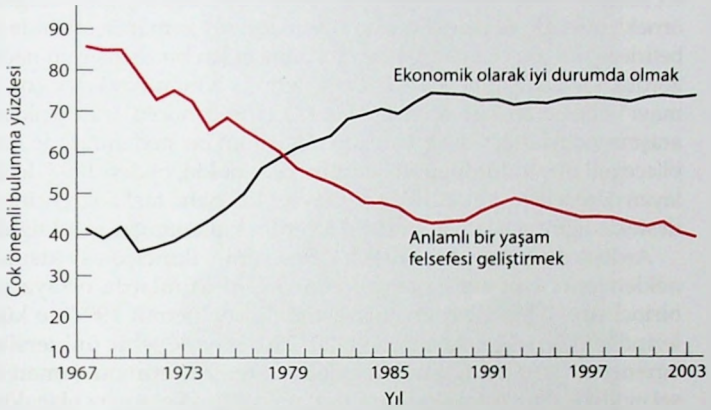
- Ardışık bağımsız örneklemler deseninde evrenden alınan farklı katılımcı örneklemleri, anketi belirli aralıklarla cevaplarlar.
- Ardışık bağımsız örneklemler deseni araştırmacılar, zaman içindeki değişimi araştırma olanağı sağlar.
- Ardışık bağımsız örneklemler deseni araştırmacıya, örneklemdeki katılımcıların bireysel olarak zaman içinde nasıl değiştiklerine dair bir çıkarım yapma olanağı vermez.
- Evrenden alınan farklı örneklemler denk olmadıkları, yani evreni eşit düzeyde temsil edemedikleri zaman, ardışık bağımsız örneklemler deseninde sorun çıkar.

Temel Kavram

Ardışık bağımsız örnekleme deseninde, zaman içerisinde (art arda gelecek şekilde) bir dizi kesitsel anket uygulaması yapılır. Örneklemler bağımsızdır, çünkü her bir zaman diliminde *farklı* bir örneklem grubu anketleri doldurur. İki önemli unsur vardır: 1) Her bir örnekleme aynı sorular sorulmalıdır ve 2) Aynı evrenden farklı örneklemler alınmalıdır. Eğer bu iki koşul sağlanırsa, araştırmacılar farklı zamanlarda alınan anket cevaplarını karşılaştırabilirler. Bu desen en çok, araştırmanın esas amacı bir evrende zamanla tutum ve davranışlarda ortaya çıkan değişimi açıklamak için uygundur. Örneğin, kamuoyu araştırmacıları sık sık ABD başkanını ne ölçüde onayladıklarını (desteklediklerini), ABD'den seçtikleri bağımsız örneklemlere sorarlar (Buna, "başkanın onay derecesi" denilir). Zaman içinde bu onay düzeyi Amerikalıların, başkanın icraatlarıyla ilgili görüşlerini nitelemek için kullanılır.

Diğer bir örnekte, 1966'dan beri her yıl yaklaşık 700 yüksekokul ve üniversiteyi temsil eden 350.000 tam zamanlı birinci sınıf öğrencisini kapsayan bir örneklem araştırılmıştır (Pryor, Hurtado, DeAngelo, Patuki Blake ve Tran, 2009; Sax ve ark., 2003). Bu araştırma projesi, yaklaşık 50 yıllık çalışma süresince 1500'ün üzerinde üniversite ve 10 milyonun üzerinde öğrencinin katılımıyla ABD'de yüksek öğretimin araştırıldığı en uzun ve kapsamlı görgül (emprical) çalışmayı temsil etmektedir. Öğrencilere birkaç konuda yaklaşık 40 soru sorulmaktadır ve on yıllar boyunca bazı sorular değişse de birçok soru her yıl sorulmaya devam etmektedir. Bu durum, bu araştırmayı bağımsız örneklemler deseninin mükemmel bir örneği yapmaktadır.

ŞEKİL 5.3 1966-2003 yılları arasında üniversite birinci sınıf öğrencilerinin değerlerinde karşıt eğilimler
Kaynak: Sax ve ark. (2003), Şekil 7 (s. 7)



Bu zaman aralığında öğrencilerin değer ve amaçlarındaki değişimlerle ilgili ne söylenebilir? Sax ve ark. (2003) öğrencilerin yaşamlarındaki anlam ve amaç ihtiyacını ölçmek için, onların çeşitli değerlere ne kadar önem verdiklerini belirtmelerini istediler. İki değerle özellikle ilgilendiler: "Anlamlı bir yaşam felsefesi geliştirmenin önemi" ve" Ekonomik olarak iyi durumda olmanın önemi". Şekil 5.3 bu değerleri "çok önemli" veya "gerekli/temel" olduğunu onaylayan öğrencilerin oranlarını gösteriyor. 1960'ların sonlarında öğrencilerin %80'inden fazlası anlamlı bir hayat felsefesi geliştirmenin çok önemli veya gerekli olduğunu belirtmiştir –aslında bu öğrenciler tarafından uygun bulunan en yüksek değerdir. Tersine, 1960'ların sonlarında ekonomik olarak iyi durumda olmak öğrencilerin %45'inden azı içi çok önemli veya gerekli olmuş ve öğrenci değerleri arasında beşinci veya altıncı olarak sıralanmıştır. Şekil 5.3'te görüldüğü gibi, değerlerdeki bu çelişen eğilimler 1970'lerin başlarında değişmeye başlamış, 1977'de çaprazlanmış ve 1980'lerin sonlarında tamamen ters dönmüştür.

2012'de yeni gelen öğrencilerin %81'i "ekonomik açıdan çok iyi durumda olma"yı kişisel bir amaç olarak "çok önemli" veya "önemli" düzeyinde sınıflamış, böylece en yüksek seviyesine çıkmıştır (Pryor ve ark., 2012). Bu bulgunun 2008'lerde başlayan ekonomik gerileme dönemi ile ilgili olması mümkündür (Pryor ve ark., 2009). Böyle bir yorum, 2012'de öğrencilerin yüksekokula katılma maliyetinden her zamankinden daha çok etkilendiklerini bildirmeleriyle de uyumludur. İlginç şekilde, 2012 yılında yeni bir soru eklenerek, ne zaman mezun olmayı umdukları sorulmuştur. Yeni gelen birinci sınıf öğrencilerinden % 84,3'ü dört yılda mezun olacaklarına inanmaktadır. Gerçekte bu oranın ancak yarısı dört yılda mezun olacaktır (bk. Pryor ve ark., 2012).

Ardışık bağımsız örneklemler deseninin kısıtlılıkları vardır. Şimdi de üniversite birinci sınıf öğrencilerin uygulanan 2012 anketinin (Pryor ve ark., 2012) başka bir bulgusunu ele alalım. Anket sosyal ve politik görüşlerle ilgili 10 soru sormaktadır. Bu sorulardan en yüksek kabul alanı, eşcinsel evliliğinin yasal hâle gelmesi

olmuştur (%75) ve 1997'de soru ilk kez sorulduğunda alınan %50.9 oranından yüksektir. 1997 ve 2012 arasındaki bu tutum değişimini nasıl açıklayabiliriz? Cevap, en azından sadece bu veriye dayanarak bunu açıklayamayacağımızdır. Bağımsız örneklemler deseninde, evrenin özelliklerinin zaman içerisinde nasıl değiştiğini belirlenebilir ama zaman içinde meydana gelen bir değişimin nedenini belirlemek zordur. Dikkat edilmelidir ki öğrencilerin "ekonomik olarak çok iyi durumda olmayı" önemli bir kişisel amaç olarak seçmelerindeki artışı açıklamaya çalışırken, araştırmadaki diğer cevapların, bu değişimin bir nedeninin de kötü ekonomi olabileceğini düşündürdüğünü belirttik. Esas nokta, neden 1997'de üniversiteye başlayan öğrencilere göre 2012'de nerdeyse %25 daha fazla öğrencinin eşcinsel evliliği desteklediğini açıklamak için başka veriler kullanılması gerektiğidir.

Ardışık bağımsız örneklemler deseninin ikinci olası kısıtlılığı, ardışık örneklemlerin aynı evreni temsil etmediği durumlarda ortaya çıkar. Üniversite birinci sınıf öğrencileri araştırmasında, örneklemin 1997'de küçük taşra kolejlerinden gelen öğrencilerden ve 2012'de büyük şehir üniversitelerinden gelen öğrencilerden oluştuğunu farz edelim. Öğrencilerin bu zaman içerisinde eşcinsel evlilikle ilgili tutumlarının karşılaştırılması anlamsız olacaktır. Yasal eşcinsel evliliğin desteklenmesi bu tür farklı okullara devam eden öğrenciler arasında farklılaşabilir ve bu durum farkı açıklayabilir. Bu farazi örnek, *denk olmayan ardışık örneklemler sorununu* göstermektedir. *Evrenin zaman içerisindeki değişimi, ancak ardışık örneklemler aynı evreni temsil ediyorsa, doğru şekilde açıklanabilir.* Denk olmayan ardışık örneklemler ile ilgili sorunlar için en iyi çözüm, çözmemize yardımcı olacak ileri istatistik yöntemleri olmakla birlikte, aynı evreni temsil eden ardışık örneklemleri dikkatlice seçerek bu sorunu en baştan önlemektir.

Boylamsal Desen

- Boylamsal desende aynı katılımcılardaki değişimi gözlemek için, aynı katılımcılar zaman boyunca incelenir.
- Anket verilerinin korelasyonel doğası gereği, geçen zaman boyunca bireylerdeki değişime neyin neden olduğunu belirlemek zordur.
- İnsanlar zaman içerisinde araştırmayı bıraktıklarında, son örneklem artık başlangıçtaki örneklem ile karşılaştırılamıyor veya evreni temsil edemiyor olabilir.

Temel Kavram

Boylamsal desenin ayırt edici özelliği aynı katılımcı örnekleminin birden fazla kez incelenmesidir. Boylamsal desenin iki önemli avantajı vardır. İlki, araştırmacı bireysel katılımcı için değişimin derecesi ve yönünü belirleyebilir. Ayrıca, her bireyin cevaplarındaki değişim ölçüldüğünden, tutum ve davranış değişiminin nedenlerini belirlemek daha kolaydır. İkincisi, araştırmacı doğal şekilde ortaya çıkan bazı olayların etkisini ölçmek istediği zaman, boylamsal desen en iyi araştırma desenidir.

Üniversiteden erken yetişkinliğe ve erken yetişkinlikten orta yaşa geçişte yeme ile ilgili tutum ve davranışların değişiminin incelenmesinde boylamsal desen kullanılmıştır (Heatherton, Mahamedi, Striepe, Fieldve Keel,1997; Keel, Baxter, Heatherton ve Joiner, 2007). Ergen ve üniversite öğrencilerindeki yeme bozuklukları hakkında çok şey biliniyor ama bireyler kariyer yapıp, hayata başlayıp, evlenip, çocuk yetiştirip daha güçlü kimlik duygusu kazanırken,

ŞEKİL 5.4 Bazı anketler (örneğin, Heatherton, Keel ve arkadaşları, 1997; 2007 gibi) insanlar büyüdükçe yeme bozukluklarından nasıl etkilendiklerini araştırmaktadır.



yeme bozukluğunun nasıl geliştiği konusundaki bilgi daha azdır. Bu araştırmacılar şöyle bir hipotez kurdular: Bireyler yetişkinlik boyunca rollerini ve amaçlarını değiştirdikçe, fiziksel görünüşe verdikleri önem azalacak ve bu, yemeyle ilgili kusurlu tutum ve davranışların yayılması olasılığını azaltacaktır (bk. Şekil 5.4).

Araştırmanın ilk oturumunda 1982'de özel bir kuzeydoğu üniversitesinden seçkisiz seçilmiş 800 kadın ve 400 erkekten oluşan bir örneklem, yeme ve diyetle ilgili bir anketi cevaplamıştır. Cevap oranı kadınlar için %78 (N = 625), erkekler için %69 olmuştur (N = 276). 1992'de araştırmacılar aynı örneklemle iletişime geçmiş (mezunlar derneği aracılığıyla) ve onlara yeme tutum ve davranışlarıyla ilgili aynı anketi tekrar vermişlerdir. Üçüncü oturumdaki veriler, 2002'de aynı bireyler kırklarının başlarındayken toplanmıştır. Boylamsal desenler zahmetlidir, ama zaman içinde bireylerde meydana gelen değişimleri inceleme olanağı sağlaması, araştırmacıları motive etmektedir .

Yeme tutum ve davranışları zaman içerisinde değişmiştir. Üniversiteden sonraki on yılda, kadınlardaki yeme bozukluğu semptomları, kronik diyet yapma ve vücut memnuniyetsizliği azalmıştır (Heatherton ve ark., 1997). Bununla birlikte, bu azalmaya rağmen kadınların kendi vücutlarıyla ilgili memnuniyetsizlikleri ve kilo verme isteklerinin yüksek olduğu gözlenmiştir. Erkekler ise aksine, üniversite sırasında yeme ve kilo ile ilgili nadiren sorun yaşamıştır. On yıl sonra kilo artışı yaşamışlardır (kadınların aldıkları ortalama 1.8 kg'a karşılık, erkekler ortalama 5.4 kg almıştır). Erkekler üniversiteyi bitirdikten 10 yıl sonra diyet ve yeme bozukluğu belirtilerinde artış bildirmişlerdir ama bu bildirim yine de kadınlara kıyasla daha düşüktür.

16

Araştırmacılar, boylamsal araştırmaları anlayışımızla ilgili bazı ilginç gözlemler yapmışlardır. Kadınlardaki yeme sorunlarının azalmasının başlıca nedenlerinin, 20'li yaşlarındaki olgunlaşma, rol değişimleri ve üniversite dışına çıkmaları (kampüste ortaya çıkan ince olma baskısını ortadan kalkması) olduğunu ileri sürmüşlerdir. Yine de katılımcılarda gözlenmiş olan değişimi, başka süreçler de açıklayabilir. Araştırmacılar farklı üniversitelerden alınan bağımsız ardışık örneklemler deseni kullanarak, yeme bozukluğu semptomlarının ve beden memnuniyetsizliğinin 1992 örnekleminde, 1982 örneklemine göre, daha düşük olduğunu bulmuşlardır (Heatherton, Nichols, Mahamedi ve Keel, 1995). Bu bulgular yeme bozukluğu tutum ve davranışlarındaki azalmanın, toplumsal düzeyde 10 yıllık değişimi (örneğin, medyadaki yeme bozukluğu ile ilgili artan bilgilerin etkisini) yansıtabileceğini düşündürmektedir. Boylamsal desenlerdeki olası bir problem, zamanla bireylerde gözlenen değişimin kesin nedenlerini tam olarak belirlemenin zor olmasıdır[1]

Üniversite mezuniyetinden 20 yıl sonraki yeme davranış ve tutumlarıyla ilgili ne söylenebilir? Yirmi yıllık araştırma boyunca genel olarak kadınlar, erkeklere göre daha fazla kilo memnuniyetsizliği, diyet ve yeme bozukluğu tutumu göstermiştir (Keel ve ark., 2007). 2002 yılındaki araştırmada araştırmacılar, vücut ağırlığının ortalama olarak hem kadın (üniversiteden sonra 6 kilo) hem erkeklerde (üniversiteden sonra 7 kilo) arttığını gözlemişlerdir. Erkeklerde, kilo almaya paralel olarak, diyet yapma ve vücut memnuniyetsizliği 2002 yılında en üst düzeydedir. İlginç şekilde, araştırmadaki kadınlar kırklarının başlarına geldiklerinde, daha az diyet, daha az yeme bozukluğu ve daha az vücut memnuniyetsizliği bildirmişlerdir. Aslında kadınlarda en fazla vücut memnuniyetsizliği üniversite yıllarında görülmüştür. İstatistiksel analizlere dayanarak araştırmacılar evlilik, ebeveynlik, kariyer aracılığıyla elde edilen yetişkinlik rollerinin, kadınlardaki yeme bozukluğunun azalmasına neden olduğunu öne sürmüşlerdir. Başka bir deyişle, üniversite yıllarında fiziksel görünüm önemliyken, evlilik ve anne olmakla ilgili öncelik değişimi kadınların incelik isteğini daha önemsiz kılmıştır.

Boylamsal desenle ilgili olası diğer bir problem, boylamsal çalışmanın örneklemine sonradan ulaşmanın zor olabilmesidir. Boylamsal desende aynı insan tekrar tekrar test edildiği (ve elbette böylece örneklem her seferinde aynı evreni temsil ettiği) için, boylamsal desenin denk örneklem sorununu çözdüğünü düşünebilirsiniz. Ne yazık ki ancak sadece ilk araştırmadaki örneklemin tüm üyeleri sonraki araştırmalara katıldığında, boylamsal desendeki örneklemler zaman içerisinde aynı kalır. Bu mümkün değildir. Örneğin, yukarıda ele alınmış olan 1997 tarihli araştırmada, 1992 yılındaki uygulamaya, 1982 örneklemindeki 901 katılımcının sadece 724'ü (%80) katılmıştır. 2002'deki üçüncü uygulamada, 1982'deki 900 katılımcının 654'ü (%73) katılmış, bunlardan 561'i

[1] Heathherton ve ark. (1997) ardışık bağımsız örneklem anketlerine kıyasla boylamsal anketlerde bireylerin yaygın yeme bozukluğundaki düşüşün bireylerin olgunlaşmalarına ek olarak toplumsal değişikliklerden dolayı olduğuna işaret etmişlerdir.

15

(%86) 1992'deki anketi de cevaplamıştır. Böylece 20 yılın sonunda araştırmacılar her üç uygulama için, (1982, 1992, 2002) başlangıçtaki 900 katılımcının %62.3'ü için cevap elde etmişlerdir.

Başlangıçtaki örneklemin tüm katılımcıları boylamsal desenin tüm oturumlarını tamamlayamazsa, *kayıp* nedeniyle muhtemel bir sorun vardır. Kayıp, muhtemelen boylamsal desenin en büyük dezavantajıdır, çünkü katılımcı sayısı zaman içerisinde azaldıkça, örneklemin alındığı evreni daha az temsil etmesi olasıdır. Yine de boylamsal desende son örneklemin ilk örnekleme denk olup olmadığını belirlemek genelde mümkündür. Başlangıçtaki örnekleme katıldıkları için takip eden oturumlarda cevap vermeyen katılımcıların özellikleri bellidir. Bu nedenle araştırmacılar, cevap vermeyen katılımcıların, devam eden katılımcılardan hangi yönlerden farklı olduğuna bakabilir.

Araştırmacılar başlangıçtaki 1982 anketine katılan ama devam etmeyenlerin cevapları ile, araştırmaya 2002 anketine kadar devam edenlerin cevaplarını karşılaştırarak kayıp ile ilgili problemleri araştırmışlardır (Keel ve ark., 2007). Araştırmaya devam etmeyenlerle karşılaştırıldığında, araştırmaya katılmaya devam edenler kendilerini daha kilolu, daha sık diyet yapan, zayıflık için daha istekli olarak tanımlamışlardır. Bu bir tepki yanlılığını gösterebilir, çünkü 2002'de devam eden katılımın nedeninin, araştırma konusuna duyulan ilgi olabileceğini düşündürüyor. Mümkündür ki 2002'deki araştırmada kilo ve vücut kaygıları, bu olası tepki oranı yanlılığı nedeniyle artmış olabilir.

Boylamsal desen bireysel değişimin belirlenmesi avantajı, aynı birey birden fazla incelendiği için ortaya çıkar. Çelişkili biçimde, boylamsal desendeki problemler de bu nedenle ortaya çıkabilir. Katılımcıların, zaman içinde değişmiş görünmektense, tutarlı olmaya çalışması da mümkündür ve bu bir problemdir. Diğer olası bir problem ise ilk anketin, katılımcıları araştırılmakta olan konuya duyarlılaştırabilmesidir. Örneğin, öğrencilerin kampüsteki suçlarla ilgili endişelerini ölçmek için kullanılan bir boylamsal deseni göz önüne alalım. Çalışma başladıktan sonra, katılımcılar suç raporlarına normalden daha fazla dikkat edebilirler. Bunun tepkisel ölçüm örneği olduğu –araştırmada yer aldıklarını bildikleri için olağandan farklı davranma– açıktır.

Boylamsal çalışmadaki katılımcılar, zaman içinde yeme tutum ve davranışlarında tutarlı olmaya çalışmaktan ziyade, üniversite yıllarındaki yeme sorunlarının aynen devam ettiğini bildirmek istememiş olabilirler. Böylece araştırmacıların 10 yıllık zaman diliminde yeme probleminde gözledikleri azalmanın nedeni, "30'una yaklaşan kadınların tipik olarak ergenlikle ilişkili olan problemler yaşadıklarını kabul etmekten utanmış olmaları" olabilir (Heatherton ve ark., 1997, s. 124). Katılımcılardan kendi tutum ve davranışlarını bildirmeleri istendiğinde, katılımcıların bildirimlerinin, onların gerçek davranışlarıyla niçin örtüşmeyebileceğini açıklama konusunda araştırmacılar dikkatli olmalıdır. Bu önemli konuya bu bölümde daha sonra döneceğiz.

ANKET FORMU

Katılımcılar örneklemi tam olarak temsil etse, tepki oranı %100 olsa ve araştırma deseni incelikle planlanmış ve mükemmel şekilde uygulanmış olsa bile, eğer anket formu kötü yapılandırılmışsa anketin sonuçları faydalı olmayacaktır. Bu bölümde, anket formunu tanıtacağız. Yararlı olabilmesi için anket formları, öz bildirim ölçeklerindeki demografik değişkenlerin ve bireysel farklılıkların geçerli ve güvenilir bir ölçümünü vermelidir. İyi bir anket formu hazırlamak gerektiği zaman, deneyim kadar değerli olmasa da, bir anket formu yapılandırma konusunda bilinmesi gereken birkaç genel ilke vardır. Anket formu hazırlamanın altı temel aşamasını anlatacağız ve ardından soruları yazma ve uygulama için özel rehber vereceğiz.

Bir Araç Olarak Anket Formları

- Birçok anket araştırması, değişkenleri ölçmek için anket formu kullanır.
- Demografik değişkenler, araştırılan insanların kişisel özelliklerini belirtir.
- Anket formunun doğruluk ve dakikliği, hazırlanırken sağlanan uzmanlık desteğine ve özene bağlıdır.
- Öz bildirim ölçekleri, kişilerin tercih veya tutumlarını ölçmek için kullanılır.

Temel Kavram

Anket araştırmasının (ve herhangi bir araştırmanın) değeri, en sonunda araştırmacının yaptığı ölçümün niteliğine bağlıdır. Bu ölçümün niteliği ise ölçüm yaparken kullanılan araçların niteliğine bağlıdır. **Anket** araştırmasında kullanılan ana araştırma aracı anket formudur. Yüzeysel olarak, çok modern bilimsel araştırmalar için anket formu kullanımı yüksek-teknolojik bir araç gibi görülmeyebilir ama uygun şekilde yapılandırılıp kullanıldığında, farklı değişkenlerin ölçümü için güçlü bir bilimsel araçtır.

Demografik Değişkenler Demografik değişkenler, anket araştırmasında sıklıkla ölçülen önemli bir değişken türüdür. Demografik değişkenler, araştırılan insanların özelliklerini belirlemek için kullanılır. Irk, etnik köken, yaş, sosyoekonomik statü gibi ölçümler, demografik değişkenlere örnek verilebilir. Araştırmacıların bu değişkenleri ölçmeye karar vermesi, araştırmanın amacının yanı sıra, bazı değerlendirmelere dayanır. Katılımcılardan kendi ırk ve etnik kökenlerini tanımlamalarını isteyerek, araştırmacılar örneklemdeki çeşitliliği belgeleyebilir ve eğer araştırma sorusuyla ilgiliyse ırk ve etnik kökene göre grupları karşılaştırabilir.

Irk gibi bir demografik değişkeni ölçmek ilk bakışta kolay görünebilir. Doğrudan bir yöntem, katılımcılara açık uçlu bir soruda basitçe ırklarını belirtmeleri istenir: Hangi ırktansınız? ..... Böyle bir yöntem dürüst olabilir ama ortaya çıkan ırk ölçümü tatmin edici olmayabilir. Örneğin, bazı katılımcılar yanlışlıkla "ırk" ve "etnik köken"i karıştırabilir. Katılımcılar ve araştırmacılar, etnik grupları belirlemede önemli olan ayrımları gözden kaçırabilirler. Örneğin, Hispanik

ŞEKİL 5.5 Etnik köken önemli bir demografik değişken olmakla birlikte, insanları bu değişkene göre doğru şekilde sınıflamak kolay bir iş değildir.



sözcüğü ırk tanımlamaz; Hispanik sözcüğü, kökeni İspanyolca konuşulan ülkelerden gelenlerin tümünün ortak ismidir. Bu yüzden İspanya'da doğan biri Hispanik olarak sınıflandırılabilir. Latin, bazen Hispanik alternatifi olarak kullanılabilen bir terimdir fakat Latin kökeni, Kanada ve ABD haricindeki Kuzey ve Güney Amerika ülkelerinden gelen insanlar için kullanılır. Bu gibi ayrımlar karıştırılabilir (bakınız Şekil 5.5). Örneğin, araştırmacıların tanıdığı birisi İspanyol kökenli bir Avrupalıdır ve kendisini Latin değil, Avrupalı saymaktadır. Bu insanın etnik kökeni Hispaniktir.

Genelde, anket araştırmasındaki ölçümde "hızlı ve kirli" yaklaşımlar, analizi ve yorumlaması zor ve karmaşık veriler ortaya çıkarır. Örneğin 2010 ABD nüfus sayımında 9 milyon insan kendini çok ırklı olarak tanımlamıştır (Jones ve Bullock, 2013). Eğer araştırmacılar bunu olası bir seçenek olarak ankete dâhil etmezlerse, katılımcılardan gelen bilgiler yanlış olabilir veya soruyu cevaplamayabilirler. Entwistle ve Astone (1994) ırk ölçümünde ihtiyatlı ve etkili bir yaklaşım önermişlerdir. Bir insanın ırkını ölçmek için dokuz soru dizisi belirlemişlerdir. Örneğin, "Kendini hangi ırktan sayıyorsun?" bu sorulardan birisidir. Diğer sorular kişinin atalarının hangi ülkeden geldiği ve Latin katılımcıların Meksikalı, Porto Rikolu, Kübalı veya başka bir yerli olup olmadığını anlamayı amaçlayan sorulardır. Bu detaylı sorular, araştırmacılara ırk ve etnik kökeni daha doğru ve dakik şekilde ölçme olanağı vermektedir. Irk ve etnik kökenin ölçümü ile ilgili bu örneği, daha genel bir ilkeye örnek olması için kullandık: *Bir araştırma aracı olarak anket formunun doğruluğu ve dakikliği, hazırlanma aşamasındaki uzmanlığa ve dikkate bağlıdır.*

Tercihler ve Tutumlar Anketlerde bireylerin tercih ve tutumları sıklıkla ölçülür. Örneğin, bir pazar araştırmacısı tüketicilerin farklı kahve markası tercihleriyle ilgilenebilir veya bir siyasi grup, seçmenlerin tartışmalı ulusal meseleler konusundaki tutumlarıyla ilgileniyor olabilir. Öz bildirim ölçekleri, katılımcıların ölçekteki maddelerle ilgili yargılarını ölçmek (örneğin; boşanma, siyasi adaylar, çeşitli hayat olayları) veya ölçekte sunulan bazı boyutlarla ilgili insanlar arasındaki farkları belirlemek (örneğin, kişilik özellikleri, stres düzeyi) için yaygın olarak kullanılır. Öz bildirim ölçümleri, genelde bir form şeklinde, psikolojide en yaygın olarak kullanılan araçlar arasındadır. Önemi nedeniyle, bu ölçümlerin dikkatlice geliştirilmesi hassas bir konudur. Öz bildirim formlarıyla yapılan ölçümlerin iki önemli özelliği, tüm ölçümler için gerekli olan özelliklerdir: Geçerlik ve güvenirlik.

Öz Bildirim Ölçümlerinin Güvenirlik ve Geçerlikleri

- Güvenirlik, ölçümün tutarlılığı anlamına gelir ve çoğunlukla test–tekrar test güvenirlik yöntemi ile değerlendirilir.
- Güvenirlik, bir ölçme aracında çok sayıda benzer madde kullanılarak, bir örneklemi ikiye ayırdıktan sonra test ederek ve uygulamada standart bir işlem kullanılarak yükseltilir.
- Geçerlik, bir ölçümün doğruluğu anlamına gelir: Ölçüm, ölçülmek isteneni gerçekten ölçüyor mu?
- Yapı geçerliği, bir ölçümün ölçülmesi planlanan teorik kavramın ne kadarını ölçtüğünü gösterir; uyum geçerliliği ve ayırt edici geçerlik ile belirlenir.

Güvenilir öz-bildirim ölçümleri, güvenilir gözlemciler veya herhangi bir güvenilir ölçme aracında olduğu gibi, tutarlılık ile nitelendirilir. Güvenilir bir öz-bildirim ölçümü, her uygulandığında benzer (tutarlı) sonuçları ortaya çıkaran ölçümdür. Öz-bildirim ölçümleri, davranışlar hakkında öngörüde bulunulacağı zaman, güvenilir olmak zorundadır. Örneğin, stresle bağlantılı sağlık problemleri hakkında öngörüde bulunabilmek için, bireylerin yaşam streslerinin ölçümü güvenilir olmalıdır. Bir testin güvenirliğini belirlemenin çeşitli yöntemleri vardır. Yaygın kullanılan yöntemlerden birisi, test-tekrar test güvenirliğini hesaplamaktır. Test-tekrar test güvenirliği yönteminde, genellikle büyük bir gruba aynı soru formu iki farklı zamanda uygulanır (yani, test ve tekrar test). Soru formunun güvenilir bir ölçme sağlaması için, kişilerin iki ölçümde de birebir aynı puanı elde etmesi gerekmez; kişilerin puanlarının dağılımdaki yerinin iki ölçümde de benzer olması gerekir. Bu konumlanma tutarlılığı, örneklemdeki her bir kişinin soru formundaki iki puanı kullanılarak hesaplanan korelasyon katsayıları ile belirlenir. İstenilen test-tekrar test güvenirlik katsayısı .80 ve üzeridir ancak korelasyon katsayısının büyüklüğü, madde sayısı ve türü gibi çeşitli faktörlere dayanır.

Bir kavramı ölçmek için bir öz-bildirim ölçümünde çok sayıda madde olması, az madde olmasından daha güvenilir olacaktır. Örneğin, bir beyzbol oyuncusunun vurma becerisini tek bir kez vuruşuna dayanarak ya da bir insanın ölüm cezasına yönelik tutumunu bir anketin tek bir sorusu ile ölçmeye çalışırsak, muhtemelen güvenilir olmayan bir ölçüm yapmış oluruz. Eğer birçok gözlem yapılarak –birçok vuruş veya birçok anket sorusu gibi– sorularda davranışların

belirli bir ortalaması yakalanırsa, ölçümlerimizin güvenirliği büyük ölçüde artacaktır (Epstein, 1979). Şüphesiz araştırmacılar, madde sayısının ne çok fazla ne de çok az olmamasını da dikkat etmelidir. Bir ankette gereğinden fazla madde olursa, bu durum cevap veren kişileri yorabilir ya da özensiz cevaplar vermelerine neden olabilir.

Genel olarak, ölçülen faktör test edilen bireyler arasında büyük farklılıklar gösteriyorsa, ölçümler daha güvenilir olur. Ölçmenin amacı, çoğunlukla bireylerin ne kadar farklılaştığını belirlemektir. Birbirinden çok farklı bireyleri ayırt etmek, az farklı bireyleri ayırt etmekten daha kolaydır. Şu örneği ele alalım. Varsayalım ki futbolcuların diğer oyunculara etkili pas verebilme yeteneklerini ölçmek istiyoruz. Örneklemimize çok sayıda basketbolcuyu dâhil edersek (örneğin, profesyonel basketbolcular, amatör basketbolcular ve lise takımında oynayan basketbolcular), iyi oyuncuları kötü oyunculardan daha güvenilir bir şekilde ayırt edebiliriz. Eğer sadece profesyonel basketbolcular test edersek onları güvenilir bir şekilde ayırt edebilmek daha zor olur, çünkü hepsinin iyi futbolcular olması beklenir. Bu nedenle bir test, çok sayıda farklı kişilerden oluşan örneklemlere uygulandığı zaman daha güvenilir olur.

Güvenirliği etkileyen üçüncü ve son etken, soru formunun uygulandığı koşullardır. Soru formunun uygulandığı ortamda dikkati dağıtacak bir şey yoksa ve formun nasıl cevaplanacağını belirten kolay anlaşılır bir yönerge varsa daha güvenilir ölçümler sağlayacaktır. Ortamda ses varken etkilendiğiniz ya da sorunun tam olarak neyi kastettiğinden emin olmadığınız durumlardaki kendi sınav performansınızı hatırlarsanız, bunu kolayca anlayabilirsiniz.

Bir anketin güvenirliğini belirlemek ve sağlamak, geçerliğine göre daha kolaydır. Geçerliğin tanımı o kadar basit ve açıktır ki, yanıltıcı olabilir – geçerli bir ölçme aracı, ölçmek istediği niteliği ölçen araçtır. Siz hiç sınavda karşılaştıkları soruların derste işlenen konularla ilgisiz olduğundan yakınan öğrencilerle karşılaştınız mı? İşte bu bir geçerlik meselesidir.

Şimdi de bir ölçümün geçerliğinin değerlendirilmesi için kullanılan birçok yöntemden biri olan yapı geçerliliğini ele alacağız. Bir ölçümün *yapı geçerliği*, ölçülmesi tasarlanan teorik yapıyı ne derece ölçtüğünü ifade eder. Bir testin yapı geçerliğini belirlemeye yönelik bir yaklaşım, başka iki tür geçerliğe dayanır: Uyum geçerliği ve ayırt edici geçerlik. Bu kavramlar, bir örnek üzerinde daha kolay anlaşılacaktır.

Psikologlar mutluluk, yaşam memnuniyeti, öz-saygı, iyimserlik ve diğer bazı değişkenler gibi iyi olma hâlinin göstergesi olan faktörleri giderek daha fazla araştırmaktadırlar (Lucas, Diener ve Suh, 1996). Yine de bu farklı değişkenlerin hepsinin aynı yapıyı mı (örneğin, iyi olma hâlini) ölçtüğü yoksa her birinin ayrı yapılar mı olduğu açık değildir. Yaptıkları çeşitli çalışmalarda araştırmacılar, katılımcılardan mutluluğun bu farklı göstergelerini ölçmek için hazırladıkları soruları cevaplamalarını istediler. Amacımıza uygun olarak, araştırmacıların üçüncü araştırmalarına ait verilerin bir kısmını ele alacağız. Bu araştırmada katılımcıların üç ölçek doldurmaları istendi: İki yaşam memnuniyeti ölçeği; Yaşam Memnuniyeti Ölçeği (Satisfaction with Life Scale - SWLS), 5 maddeli Yaşam Doyumu ölçeği (Life Satisfaction - LS-5) ve Olumlu Duygu ölçeği (Positive Affect - PA) ölçeği. Bu örnekteki meselemiz, yaşam doyumu kavramının – ani, mutluluğun niteliğinin– genel olarak mutlu olma durumundan (olumlu duygu) ayırt edilip edilemeyeceğidir.

TABLO 5.1 YAPI GEÇERLİLİĞİ ÖRNEĞİ*

	SWLS	LS-5	PA
SWLS	(.88)		
LS-5	.77	(.90)	
PA	.42	.47	(.81)

*Veri Lucas ve arkadaşları (1996) tarafından sağlanmıştır, Tablo 3.
Not: SWSL = Yaşam Memnuniyeti Ölçeği; LS-5= 5 maddeli Yaşam Memnuniyeti Ölçeği; PA= Olumlu Duygu Ölçeği

Tablo 5.1, "yaşam doyumu" yapı geçerliliğinin nasıl değerlendirildiğine dair, korelasyon matrisi şeklinde veri sunmaktadır. Korelasyon matrisi, birden fazla korelasyonu göstermenin kolay bir yoludur. Öncelikle çapraz şekilde parantez içinde gösterilen değerlere bakın. Parantez içindeki korelasyon katsayıları her bir ölçeğin güvenirlik değerlerini göstermektedir. Görüldüğü gibi, üç ölçeğin de güvenirlikleri iyi düzeydedir (her biri .80'in üzerindedir). Ancak bizim burada odaklandığımız nokta, "yaşam doyumu"nun yapı geçerliğini ölçmektir; bu yüzden, tablo 5.1'de başka ne olduğuna bakalım.

Yaşam Doyumu Ölçeğindeki (SWLS) puanların, 5 maddeli Yaşam Doyum Ölçeği puanlarıyla korelasyonlu olmasını beklemek akla yakın görünüyor; ne de olsa her iki ölçek de yaşam doyumu kavramını değerlendirmek için tasarlanmıştır. Gerçekten de, Lucas ve arkadaşları bu iki ölçüm arasında .77'lik bir korelasyon bulmuştur ki bu da beklenen bir korelasyondur. Bu sonuç, ölçümlerin *uyum geçerliği* olduğunu düşündüren bir bulgudur; yani bu iki ölçüm yaşam doyumunu ölçme konusunda birbiriyle uyumludur.

Ölçümlerin ayırt edici geçerliğinin olduğu gösterildiği zaman, yaşam doyumunun yapı geçerliği çok daha güçlü bir şekilde kabul edilir. Tablo 5.1'de görüldüğü gibi, Yaşam Doyum Ölçeği (SWLS) ile Olumlu Duygu (.42) ve 5 maddeli Yaşam Doyum Ölçeği (LS-5) ile Olumlu Duygu (.47) arasındaki korelasyonlar daha düşüktür. Bu sonuçlar, yaşam doyumu ölçümlerinin, olumlu duygu adı verilen bir diğer teorik yapının ölçümü ile ilişkisinin zayıf olduğunu gösterir. Yaşam doyumu ölçekleri ile olumlu duygu ölçeği arasındaki düşük korelasyonlar, bize farklı yapıların (kavramların) ölçüldüğünü işaret etmektedir. Olumlu duygu ile yaşam doyumunu birbirinden ayırt etmesi, yaşam doyumu ölçümlerinin ayırt edici geçerliği olduğunu gösterir –yani, bireyin kendi yaşamında duymuş olduğu doyum, genel mutluluk hâli ile aynı değildir. Yaşam doyumunun yapı geçerliği bu örneğimizde desteklenmektedir, çünkü uyum geçerliği ve ayırt edici geçerliğin her ikisine dair kanıtlar sağlamaktadır.

Bir Anket Formunun Hazırlanması

- Bir anket formunun hazırlanması için, hangi bilgilerin elde edilmesi gerektiğine, anketin nasıl uygulanacağına, önceden bir soru formu taslağının hazırlanmasına, bu formun önceden denenmesine ve uygulama için açık bir yönergenin yazılmasına karar verilmesi gerekir.
- Soru cümleleri açık, net ve basit olmalı ve bilindik kelimeler kullanılmalıdır.
- Formdaki soruların hangi sırayla sorulacağı meselesine dikkat edilmelidir çünkü sıralama deneklerin cevaplarını etkileyebilir.

Anket Formu Hazırlamanın Aşamaları Güvenilir ve geçerli ölçümler sağlayacak bir anket formu oluşturmak zorlu bir iştir. Bu kısımda özellikle bir araştırma projesinde görev alıp ilk kez bir form oluşturuyorsanız, bu zorluklarla karşılaştığınızda sizlere yardımcı olacağını düşündüğümüz bazı adımlar öneriyoruz.

Adım 1: *Hangi bilgilerin elde edilmesine gerektiğine karar verin.* Sizlerin kararı, sorulacak soruların niteliğini açık bir şekilde belirleyecektir. Önerilen soru formunun olası sonuçlarını öngörmek ve "bulguların" sizin amacınızı karşılayıp karşılamadığına karar vermek önemlidir. Kötü tasarlanmış bir soru formunun uygulanması ve analiz edilmesi çok zamanınızı alacaktır; sadece iyi planlanmış bir soru formu size yorumlanabilir bulgular sağlayacaktır.

Adım 2: *Soru formunun nasıl uygulanacağına karar verin.* Kararın belirlenmesinde, çalışmaya uygun olan anket yönteminin seçilmesi önemlidir: Posta? Telefon? İnternet? Bire bir görüşme? Bu konuda sizler, her bir yöntemin avantaj ve dezavantajlarını gözden geçirmek isteyeceksiniz.

Adım 3: *Soru formunun bir taslağını hazırlayın.* Eğer başka araştırmacılar hâlen bir ölçme aracı kullanıyorlarsa, kendi ölçme aracınızı geliştirmenin gerekmediğini aklınızda tutun. Aynı formu kullanmak, kendi sonuçlarınızı önceki çalışma sonuçları ile doğrudan karşılaştırma olanağı sunar. İhtiyacınız karşılayacak hiçbir ölçme aracı yoksa, o zaman kendi taslağınızı hazırlamanız gerekir. Cümlelerin yapısı ve soruların sıralanmasındaki kuralları dikkatlice inceleyin.

Adım 4: *Soru formunu tekrar inceleyin ve gözden geçirin.* Bu önemli bir aşamadır ve anket konusunda bilgili kişilerle birlikte, sizin odaklandığınız konuda uzman olanlar taslağı inceledikten sonra yapılmalıdır. Örneğin, öğrencilerin kampüsteki yemek servisine karşı tutumlarını ölçen bir anket yapıyorsanız, yemek servisi yöneticisinin sizin formunuzu gözden geçirmesi makul olabilir. Tartışmalı bir konu ile ilgileniyorsanız, her iki tarafın temsilcileri, sorularınızdaki olası ön yargıları size gösterebilir.

Adım 5: *Soru formunun önceden denenmesi.* Son uygulamada, çalışmanın uygulanacağı asıl katılımcılara benzer bir gruba formu uygulayın. Hem spesifik maddeler hem de formun bütünü hakkında geri bildirim isteyin. Anketi önceden deneme, (görüşmeci uygulayacaksa) anketi kurallara uygun şekilde uygulayıp uygulamadığı denetlenecek olan görüşmeci için de bir prova niteliği de taşır.

6. Aşama: *Soru formuna son hâlini verin ve uygulamada izlenecek adımları yönergelerle belirleyin.* Ön denemeden gelecek sonuçlara göre formun uygulamasında ya da cümle yapılarında karşılaşılan problemleri düzeltin.

Soruların Etkili Şekilde İfadesi İçin Rehber Hukukçular, bir sorunun nasıl cevaplanacağının, o sorunun nasıl sorulduğuna bağlı olduğunu çok eskiden beri bilmektedir. Anket çalışması yürüten araştırmacılar da bu ilkeye dikkat etmelidir. Bu husus insanların varsayımsal bir grip salgını esnasında, zor bulunan bir aşının dağıtımı hakkındaki düşüncelerini araştıran bir çalışmada gösterilmiştir (Li, Vietri, Galvani ve Chapman, 2010). Bu araştırmacılar, anketi cevaplayanların aşıların dağıtımı (daha doğrusu, kimlerin yaşayacağı ve kimlerin öleceği) hakkındaki kararlarının, aşı yapma politikalarını açıklayan metinde "hayat kurtarma" teriminin mi yoksa "insanları kaybetme" teriminin mi kullanıldığına göre değiştiğini bulmuşlardır. Yani soruların ifade edilme şekli, cevap veren

kişilerin insanların hayatlarının değeri hakkındaki kararlarını etkilemiştir. Normal bir ankette her bir soru için tek bir ifade olur, ama ne yazık ki belirli bir anketteki soruların ifade şekli, hemen hemen hiçbir zaman mükemmel olamaz.

Anket çalışması yapan araştırmacılar, genellikle iki farklı formdan birisini seçerler: Birinci tip form serbest cevaplı (açık uçlu) sorular içerir ve ikinci tip form ise kapalı (çoktan seçmeli) sorular içerir. Serbest cevaplı sorular, bir okul sınavındaki deneme yazısı soruları gibidir; cevapta neyin ele alınacağını, yaklaşık olarak belirler. Örneğin, "Kürtajın yasal olması hakkındaki görüşünüz nedir?" sorusu, serbest cevaplı bir sorudur. Tersine, kapalı sorular verilecek cevabı sınırlar. Şu soru tipi ise kapalı sorudur: "Polis korumasını nasıl değerlendiriyorsunuz? Çok iyi, iyi, normal, kötü, çok kötü."

Açık uçlu soruların birincil üstünlüğü, katılımcıya kapalı uçlu sorularda olmayan bir esneklik sağlamasıdır. Bununla birlikte bu üstünlük, serbest cevaplı soruların puanlanmasının ve kaydedilmesinin neden olduğu zorluk nedeniyle ortadan kalkar. Örneğin, serbest cevaplı sorulara verilen cevaplar içinde birbiriyle bağlantısız ögeler olabilir; bu bağlantısız unsurları özetleyebilmek için çoğunlukla geniş bir kodlama yapmak gerekir. Öte yandan, kapalı sorularda çok az puanlama problemi yaşanır ve hem hızlı hem de kolay cevaplanır. Aynı zamanda kapalı sorularda cevapları özetlemek çok daha kolaydır çünkü cevaplar katılımcılar arasında çabucak karşılaştırılabilir hâldedir. Kapalı soruların en belirgin dezavantajı, katılımcıların doğallığını ve kendilerini dışa vurma olanaklarını azaltmasıdır. Ayrıca, katılımcılar kendi bakış açılarını gerçekten ifade eden bir seçenek sunulmadığı durumlarda, daha az tercih edecekleri cevapları seçmek zorunda kalabilirler. Bu nedenle, cevaplar katılımcıların görüşlerini tam olarak yansıtmıyor olabilir.

Anket sorularını etkili şekilde ifade edebilmeniz için hazırladığımız rehber aşağıdadır.

- *Bütün katılımcıların bildikleri, dolaysız ve basit sözcükler kullanın.* Kullandığınız ifadeleri, katılımcıların kendi eğilim ve kavrayışlarına göre yorumlayacaklarını aklınızda tutun. "Az", "yakın zamanda", "genellikle" gibi sözcükler veya "küresel ısınma", gibi terimler, bireyler tarafından farklı anlaşılabilir. Katılımcılar aynı zamanda, formdaki sözcüklerin kendi kültür veya alt kültürlerindeki gibi kullanıldığını farz etme eğilimdedir. Günümüzün popüler kültüründe "kötü", "iyi" anlamına gelir mi? Katılımcılar makul bir şekilde, eğer anketör bir soru soruyorsa, bu sorunun cevaplanabilmesi gereken bir soru olduğunu düşünürler. Böyle bir durum ise katılımcıların, aslında cevabı olmayan sorulara cevap vermelerine neden olur.
- *İki anlamlı, yönlendirici ve şaşırtmacalı sorulardan kaçının.* Tablo 5.2, bu gibi sorunları olan soruları ve bunlardan nasıl kaçınılabileceğini göstermektedir.
- *Mümkün olduğu kadar kısa yazın (en çok 20 sözcük).* Kısa cümlelerin anlaşılması kolaydır ve katılımcının soru formunu doldururken hızlı ve verimli hareket etmesini sağlar. Aynı zamanda, anketinizin uzunluğunu da gözden kaçırmamalısınız. Sorular, katılımcının sıkılacağı veya çabucak cevaplayıp kurtulmayı düşüneceği kadar çok mu? Soru formuna verilecek bu tepkilerin her ikisi de elde ettiğiniz cevapların niteliğini etkileyebilir.
- *Koşullu bir ifadeyi, asıl sorudan önce kullanın.* Örneğin, "Mevcut işinizden ayrılmak zorunda kalsanız, ne tür bir iş arardınız?" diye sormak " Ne tür

TABLO 5.2 ANKET SORULARININ POTANSİYEL SORUNLARI

Sorun	Örnek	Açıklama	Çözüm
Çift Amaçlı Olma	*Geçtiğimiz hafta hiç baş ağrısı ve nezle sorunu yaşadınız mı?*	Kişi, aynı anda iki sorunu birden yaşamadığı için "hayır" demiş olabilir, veya ikisinden sadece birisini yaşamışsa "evet" demiş olabilir.	İki ayrı soru olarak yazılmalı
Yönlendirme	*Pek çok insan güneş enerjisi kullanımını destekliyor. Siz ne düşünüyorsunuz?*	Eğer katılımcı kendi gerçek fikrini göz önüne almadan, "çoğu insanın" yaptığı gibi cevap verme zorunda hissederse, tepki yanlılığı ortaya çıkar.	Katılımcıyı belirli bir cevaba doğru yönlendirmeyin. "Güneş enerjisi kullanımı hakkında ne düşünüyorsunuz?" sorusu, çok daha iyi bir sorudur.
değer yüklü	*Eyalet senatörünüzün yürüttüğü ırkçı propagandaya katılıyor musunuz ?*	Duygu-yüklü sözcük (ırkçı, propaganda) içeren sorular tepki yanlılığı oluşturabilir.	Değer yüklü terimlerden kaçının ve sorularınızı, farklı görüşleri temsil eden kimselere inceletin.

bir iş arardınız ki eğer mevcut işinizden ayrılmak zorunda kalsanız " diye sormaktan daha iyidir.

Potansiyel tepki yanlılığından kaçının. Tepki yanlılığı hangi durumlarda ortaya çıkar? Katılımcı dereceleme ölçeğinde yalnızca uç değerleri veya sadece orta dereceyi seçtiği durumlarda veyahut her maddeye katıldığı (veya katılmadığı) durumlarda ortaya çıkar. Bir kavramı değerlendirmek için birden fazla madde kullanıldığında, bazı maddelerin zıt yönlü yazılması önemlidir. Örneğin, bir duygusal sağlık değerlendirmesi şu maddelerden oluşabilir:

Duygu durumum genellikle olumludur.
1 ------ 2 ------ 3 ------ 4 ------ 5
Hiç katılmıyorum — Tamamen katılıyorum

Çoğunlukla üzgünüm.
1 ------ 2 ------ 3 ------ 4 ------ 5
Hiç katılmıyorum — Tamamen katılıyorum

Bütün maddelere katıldığını belirterek tepki yanlılığı gösteren katılımcılar, her iki ölçekte de 5'i işaretleyebilirler ve bu durum duygusal sağlık ölçümünün güvenilir olmadığını gösterir. Bu "karşıt yöndeki" maddelerin cevapları için "karşıt yönde puanlama" (1 = 5, 2 = 4, 4 = 2,5 = 1) kullanılır. Katılımcıların cevapları toplandığı zaman duygusal sağlık toplam puanı elde edilir.

- *Anket formunuzun okunabilir olup olmadığını kontrol edin.* Araştırma konunuza aşina insanlardan soru formunuzu dikkatle okumalarını ve kelime ve cümlelerin belirsiz veya kafa karıştırıcı olup olmadığını kontrol etmelerini isteyin.

Bir ankette soruların ifadesinin sorun oluşturabileceği çok açık olmakla birlikte, çözümü o kadar açık değildir. *En azından önemli soruların açık ifadesi, katılımcıların cevaplarını betimleyen verilerle birlikte daima bildirilmelidir.* Soruların ifade şeklinin potansiyel etkileri, davranış incelenirken neden çoklu yöntemlere ihtiyaç olduğunun bir başka göstergesidir.

Soruların Sırası Bir ankette soruların sıralanışı özenli bir yaklaşım gerektirir. İlk bir kaç soru, anketin geri kalanının havasını verir ve katılımcının takip eden sorular üzerinde isteyerek ve özenli bir şekilde çalışıp çalışmayacağını belirler. Öz bildirim formlarında, katılımcının dikkatini soruların üzerinde toplayacak ilgi çekici sorularla başlamak en iyisidir. Demografik veri, katılımcının kendisinin cevapladığı formlarda, en sonda alınmalıdır. Ama yüz yüze veya telefonla yapılan görüşmelerde, demografik sorular en başta sorulmalıdır, çünkü cevaplaması kolaydır ve katılımcının kendisine güven duymasını sağlar. Ayrıca, daha hassas konular hakkındaki sorularından önce, katılımcıyla dostane bir ilişki kurması için görüşmeciye fırsat sağlar.

Soruların sıralanma şeklinin önemli etkileri olabilir. Araştırmacıların kürtaj konusunda biri genel, biri spesifik iki sorunun sırasını değiştirdikleri bir araştırmayı ele alın. Genel soru, "Hamile bir kadın eğer evli ve daha fazla çocuk sahibi olmak istemiyor ise yasal kürtaj hakkı alacağını düşünür müsünüz?" idi. Spesifik soru ise, " Bebeğin sakat doğma olasılığı varsa, hamile kadınlara yasal kürtaj hakkı verilmesi gerektiğini düşünür müsünüz? " idi. Genel soru, spesifik sorudan önce sorulduğunda katılımcıların %60.7'si "evet" derken, sonra sorulduğunda katılımcıların yalnızca %48'i "evet" dedi. Spesifik soru birinci ve ikinci sıralarda sorulduğunda ise bu oranlar %84 ve %83 oldu. Bu sorunla baş etmenin genel olarak kabul edilen yolu, *huni sorular* (funnel questions) kullanmaktır. Bunun anlamı çok genel bir soruyla başlamak ve aynı konuda daha spesifik sorularla ilerlemektir.

Anket sorularının sıralanma şekline dair ele alacağımız son boyut *filtre soru* kullanımıdır –spesifik soruları cevaplayacak olanları bulmak için katılımcılara genel sorular sormak. Örneğin, "Arabanız var mı?" sorusu, arabanın sahibi olmanın maliyetleri hakkındaki bir dizi sorunun başlangıcı olabilir. Bu örnekte, katılımcı eğer genel soruya "evet" derse, spesifik sorular sorulur. Eğer cevap "hayır" ise, görüşmeci bu soruları sormaz (kendi kendine doldurulan soru formlarında, katılımcıya bu bölümü atlaması için yönerge verilir). Eğer filtre soru, nesnel bir bilgi soruyorsa (örneğin, "65 yaşın üzerinde misiniz?"), kullanımı nispeten kolaydır. Bununla birlikte, davranışlara veya tutumlara ilişkin sorular filtre soru olarak kullanıldığı zaman, uyarı yapılmalıdır. Smith (1981) ilk olarak katılımcılara "hayal edebilecekleri herhangi bir ortamda" bir diğer insana vurmayı onaylayıp onaylamadıklarını sordu. Mantıken, bu en genel soruya olumsuz cevap,

herhangi bir spesifik soruya da olumsuz cevap verilmesi gerektiğini düşündürür. Fakat bu genel soruya "hayır" diyenlerin %80'den fazlası, daha sonra kendini savunmak gibi spesifik durumlarda bir diğer insana vurmayı onayladıklarını söylediler. Böyle bulgular olmakla birlikte, filtre sorular dikkatli kullanılmalıdır. Katılımcıyı mümkün olduğu kadar az meşgul etme ihtiyacı, filtre soruları etkili bir soru formu tasarlamanın temel bir aracı hâline getirir.



İyi yürütülen bir anket, araştırmanın betimleme ve yordama amaçlarına etkili bir katkıda bulunur. Binlerce değilse bile düzinelerce bireye dağıtıldığı zaman, ortalama ebatta bir soru formu bile, binlerce cevap üretecektir. Ayrıca, gördüğümüz gibi, internet kullanımıyla araştırmacılar kısa bir süre içinde gerçekten milyonlarca cevaba ulaşabiliyorlar. Fakat bir sorun var! Bu kadar çok cevapla nasıl baş edilebilir? Cevap şudur: Dikkatli bir planlama yaparak!

Anketten elde edilen cevapların analizi, anket maddeleri yazılmadan önce ele alınmalıdır. Açık uçlu sorular kullanılacak mı? Asıl amaç betimleme midir? Örneğin, esas olarak bir evrendeki oranlar veya yüzdelerle mi ilgileniliyor? Amaç korelasyonel midir; örneğin, bir soruya verilen cevapların bir başka soruya verilen cevaplarla olan ilişkisi midir? Katılımcılar evet-hayır şeklinde mi cevap verecek yoksa evet-belki-hayır şeklinde mi? Öz bildirim ölçekleri mi kullanılacak? Bu cevap şekillerinin her birisi farklı türde veriler sağlar. Daha önce öğrendiğiniz gibi, açık uçlu cevap formlarından gelen niteliksel veriler, kodlama kurallarını ve kodlayıcılar arası güvenirliği bulma yöntemlerini kullanmayı gerektirecektir. Evet-hayır cevaplarından gelen kategorik cevaplar kategorik veri sağlarken, ölçeklerin tipik olarak eşit aralıklı veri sağladığı kabul edilir (Ölçek tipleri hakkında yorumlar için Bölüm 4'e bakınız.). Bu veri biçimleri, farklı istatistiksel analiz yaklaşımlarını gerektirir.

Tasarlanan soru formunun muhtemel sonuçlarını önceden yordamak ve daha sonra bu sonuçların araştırma sorusunu cevaplayıp cevaplamayacağına karar vermek önemlidir. Sonuçları yordadığınız zaman, bu sonuçların uygun şekilde analiz edilip edilemeyeceğinden emin olmak isteyeceksiniz. Bir diğer ifadeyle, araştırmaya başlamadan önce bir analiz planınız olmalıdır. Bu planlama aşamasında, uygun istatistiksel analizi belirlemek için deneyimli anket araştırmacılarının önerilerini almanızı öneririz.

Bir kez daha, bu kitabın 11'inci ve 12'nci bölümlerinden yararlanarak istatistik işlemlere aşinalık kazanmanızı öneriyoruz. Eğer sizi, kategorik (nominal) değişkenler arasındaki ilişkileri (korelasyonları) araştırmaya götürecek olan bir anket araştırması ile ilgileniyorsanız, bu kitapta öğreneceklerinizden daha fazlasına ihtiyacınız olacaktır. *Nominal değişkenler arasındaki İlişkileri incelemek için uygun istatistik analizi, ki-kare testidir.* Bu analiz bütün istatistiğe giriş kitaplarında vardır (örneğin, Zechmeister ve Posavac, 2003). Eğer eşit aralıklı ölçek verileri ile korelasyon yapacaksınız, Pearson korelasyon katsayısı uygundur. Bu analiz tipinden Bölüm 4'te gözlemciler arası güvenirlik anlatılırken söz eettik. Bu bölümün sonuna doğru, korelasyonel analizlerden biraz daha bahsedeceğiz.

ANKET ARAŞTIRMALARI ÜZERİNE ELEŞTİREL DÜŞÜNME

Bildirilen Davranış ile Gerçek Davranış Arasındaki Uyum

- Bir anket araştırması, tepkisellik içeren bir ölçümdür, çünkü bireyler tepkilerinin kaydedildiğinin farkındadırlar.
- Sosyal beğenilirlik, katılımcıların bazen gerçekten nasıl davrandıklarından çok nasıl davranmaları "gerektiğine" dair cevaplar vermelerine neden olabilecek bir baskıdır.
- Araştırmacılar anketle elde ettikleri cevapların doğruluğunu, bu cevapları daha önceki verilerle ya da davranış gözlemleriyle karşılaştırarak değerlendirebilirler.

Anket verileri ne kadar dikkatli toplanmış ve analiz edilmiş olursa olsun, bu verilerin değeri anket sorularına cevap verenlerin dürüstlüklerine bağlıdır. Anketlere verilen cevapların insanların düşünce, görüş, duygu ve davranışlarını doğru yansıttığına inanmalı mıyız? Sözel beyanların güvenirliği meselesi, uzun zamandır tartışılmakta olduğu hâlde henüz açık bir sonuca ulaşılamamıştır. Bununla birlikte, bizler gündelik hayatta başkalarının sözel beyanlarını genellikle geçerli kabul ederiz. Bir arkadaşınız okuduğu romandan hoşlandığını söylerse bunun sebebini sorabiliriz fakat bu ifadenin arkadaşınızın duygularını gerçekten de yansıtıp yansıtmadığını genellikle sorgulamayız. Ancak gündelik hayattaki bazı durumlarda birinin ifadesinin dürüstlüğünden şüphelendiğimiz zamanlar da olur. Örneğin ikinci el araba alacağımız zaman satıcının ifadelerine her zaman güvenmek istemeyebiliriz. Ama eğer elimizde şüphelenmemizi gerektirecek bir kanıtımız yoksa, insanların söylediklerine genellikle inanırız. Anketlerden elde ettiğimiz cevaplar için de aynı durum geçerlidir.

Eşyanın tabiatı gereği, anket araştırmaları tepkisellik içerir. Cevap verenler, cevaplarının kaydedildiğini bilirler ve bu cevaplarının bazı sosyal, politik veya ticarî eylemleri etkileyebileceğini de düşünebilirler. Bu düşünce katılımcıları, gerçekte nasıl davrandıklarını değil, nasıl davranmaları "gerektiğini" bildirmeye yönelten bir baskıya dönüşür. Bu baskıları belirlemek amacıyla **sosyal beğenilirlik** kavramı sıkça kullanılır ("politik doğru" terimi de aynı baskıları işaret eder). Örneğin muhtaç birine yardım etmeye istekli olup olmadıkları sorulduğunda, katılımcılar "evet" diyeceklerdir. Çünkü bunun sosyal olarak en kabul edilir tutum olduğuna inanırlar. Gözlemsel araştırmalarla doğrulandığı gibi, anket araştırmalarında tepkisel ölçüme karşı en iyi korunma, ölçüm sırasında katılımcıların tepkisel olduklarının farkında olmaktır.

Temel Kavram

Bölüm 4'te bireylerin öz bildirimlerinin onların sübjektif yorumlardan nasıl etkilendiğini ve davranışlarını etkileyen faktörlerin nasıl farkında olmadıklarını anlatmıştık (Kutu 4.1'e bakınız). Amerikalı ve Meksikalı öğrencilerin sosyal girişkenliklerine dair kendi bildirimleri, aynı özelliğin davranışsal ölçümleriyle uyumlu çıkmamıştı (Ramirez-Esparza ve ark., 2009; Tablo 4.4'e bakınız). Dolayısıyla öz bildirime dayanan kişilik değişkeni ölçümleriyle davranışsal ölçümler birbiriyle uyumsuz çıkabilmektedir.

Hatta, insanlara verilen varsayımsal bir durumda yapacaklarını söyledikleri davranış ile o durum gerçekleştiğinde yaptıkları davranış arasında fark olduğu, birçok araştırmada gösterilmiştir (bakınız, Baumeister ve ark., 2007). Artık klâsik olmuş bir araştırmada (Latane ve Darley, 1970) insanların tek başınayken, etrafta başkalarının da olduğu duruma kıyasla, bir mağdura daha fazla yardımcı

ŞEKİL 5.6 İnsanlar bu gibi durumlarda yapacaklarını söyledikleri davranışlarla gerçekte yapacakları davranışlar farklı olabilir.



oldukları bulunmuştur. Ama ikinci bir katılımcı grubuna, çevrelerinde başkalarının olmasının, onların bir mağdura yardım davranışını etkileyip etkilemeyeceği sorulduğu zaman, hepsi de etkilenmeyeceğini söylemişlerdir. Dolayısıyla bireylerin sözlü bildirimleri ile gerçek davranışları *örtüşmeyebilir* (bk. Şekil 5.6). Bu gibi araştırma sonuçları, davranışlar hakkında sadece sözel bildirimlere dayanarak elde edilen sonuçlara çok daha temkinli yaklaşmamız gerektiğini gösterir.

Gözlenen davranış ile sözlü bildirim arasındaki potansiyel çelişki, bir defa daha, çok yöntemli yaklaşımın davranışı ve zihinsel süreçleri anlamamıza yardımcı olma yeteneğini gösteriyor. Örneğin, sözlü bildirimlerin doğruluğu bazen arşiv kayıtlarına başvurularak kontrol edilebilir. Bildirilen yaş, oy verme davranışı, ev sahipliği, yardımseverliğin düzeyi ve bir ankete verilen diğer tepkiler, belki de mevcut kayıtlar incelenerek araştırılabilir (örneğin, Parry & Crossley, 1950). Sözlü bildirimlerin doğruluğu, mümkün olan her durumda, davranış doğrudan gözlenerek değerlendirilebilir ve değerlendirilmelidir (örneğin, Parry & Crossley, 1950).

İlişkisellik (Korelasyon) ve Nedensellik

- İki değişken ilişkili olduğunda, bu değişkenler için kestirimlerde bulunabiliriz. Ama aralarında sadece korelasyon olduğu bilgisine dayanarak, bu ilişkinin nedenini belirleyemeyiz.
- İki değişken arasındaki ilişki üçüncü bir değişken tarafından açıklanabiliyorsa, bu ilişkiye "sahte ilişki" denilir.
- Çok yöntemli yaklaşımlarla birlikte ilişkisel kanıtlar, bir davranışın potansiyel nedenlerini belirlemede araştırmacılara yardımcı olabilirler.

Anketler genellikle ilişkisel araştırmalarda kullanılır ve ilişkisel araştırma, betimleme ve yordama gibi bilimsel amaçlara ulaşmak için mükemmel bir yöntemdir. Örneğin, fiziksel sağlıkla psikolojik iyi oluş (well-being) arasındaki ilişkileri araştıran çalışmalar, araştırmacıların sağlıkla sağlık sorunları hakkında yordama yapmalarına olanak sağlarlar.

İlişkisel kanıt, araştırmacıların ilişkili değişkenler hakkında öngörüde bulunmalarını sağlar. Fakat çok bilinen "ilişki, nedeni göstermez" ifadesi, sadece iki değişken arasındaki ilişkiye dayanarak nedensel çıkarımlar yapma

olanağımızın çok kısıtlı olduğunu bize hatırlatır. Örneğin, hayattan memnun oluş ile dışarı çıkma (sosyal aktivite) arasında güçlü bir ilişki vardır (Myers ve Diener, 1995). Yine de sadece bu ilişkiye bakarak, daha fazla dışarı çıkıp daha çok sosyal aktiviteye katılmanın hayattan daha fazla memnun olmaya neden olacağını inandırıcı şekilde iddia edemeyiz. Dışarı çıkmanın hayattan daha çok memnun olmaya neden olma olasılığı kadar "tersi" bir nedensel ilişki de doğru olabilir: Hayatından memnun olma, insanların daha fazla dışarı çıkıp daha sosyal olmalarına da neden olabilir. Sadece bu iki değişken arasındaki ilişkiyi bilerek doğru bir nedensel yön tayin etmek olanaksızdır.

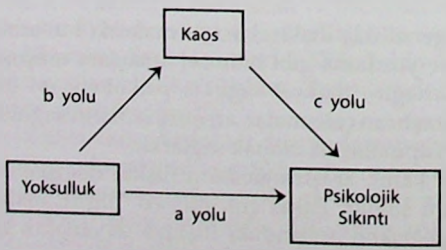
İlişkisel analizlerde ilişkinin yönünü belirleyememek, karşılaştığımız zorluklardan sadece birisidir. İki değişken arasındaki ilişki için başka bir nedensel açıklama da mümkündür. Örneğin, arkadaş sayısı gibi üçüncü bir değişken, insanların daha fazla dışarı çıkmalarına ve hayatlarından daha memnun olmalarına neden olabilir. Üçüncü bir değişkenle açıklanabilen bir ilişkiye **sahte ilişki** denir (Kenny, 1979). Bu örnekte, birisinin kendi hayatından memnun olması ile dışarı çıkma sıklığı arasındaki ilişkiyi açıklayabilecek "arkadaş sayısı" gibi üçüncü bir değişken olabilir. Daha fazla arkadaşı olanlar, daha az olanlardan daha fazla dışarı çıkıyor ve hayatlarından daha fazla memnun oluyor olabilir. Bu, yaşam doyumu ile dışarı çıkma arasında gerçek bir pozitif ilişki olmadığı anlamına gelmez (elbette vardır); bu sadece, bu ilişkinin neden ortaya çıktığını açıklayabilecek başka değişkenlerin (örneğin, arkadaş sayısı) ölçülmediğini gösterir.

Temel Kavram

Sadece iki değişken arasındaki ilişkiye bakılarak neden nedensel çıkarım yapılamayacağını anlamak çok önemlidir. Bununla birlikte, bu ilişkisel kanıtın, davranışın olası nedenlerini belirlemede çok kullanışlı olduğunu bilmek de aynı derecede önemlidir. İz analizi gibi ileri istatistik teknikleri, ilişkisel çalışmaların nedensel yorumlanmasına yardımcı olabilir (Baron ve Kenny, 1986; Holmbeck, 1997). İz analizi, aracı (mediator) ve yönlendirici (moderator) değişkenlerin belirlenmesini içerir. Aracı değişken, iki değişken arasındaki ilişkiyi açıklamakta kullanılan değişkendir. Yönlendirici değişkense iki değişken arasındaki ilişkinin yönünü ve gücünü etkileyen değişkendir.

Şekil 5.7, çocukların psikolojik uyumları üzerinde yoksulluğun etkilerinin araştırıldığı bir çalışmada, yönlendirici değişkenin bir örneğini göstermektedir (Evans, Gonnella, Marcynyszyn, Gentile ve Salpekar, 2005). Bu araştırmacılar, önceki araştırmalarla uyumlu şekilde, psikolojik sıkıntı ile yoksulluk ölçümleri

ŞEKİL 5.7 Yönetici değişken örneği



arasında bir ilişki gözlemişlerdir: Fazla yoksulluk, çocuklarda fazla sıkıntı demektir (Şekil 5.7'deki a yolu). Araştırmacılar ayrıca, bu ilişkiyi açıklamak için kaosu yönlendirici bir değişken olarak öne sürmüşlerdir. Onlara göre kaotik yaşama koşullarının özellikleri belirsizlik, karışıklık, bina yokluğu, gürültü, aşırı kalabalık ve kalitesiz konutlardır ve bu yaşama koşulları, yoksulluk ile çocukların psikolojik sıkıntıları arasındaki ilişkiyi açıklayabilir. Bu da Şekil 5.7'deki b ve c yollarıyla gösterilmiştir.

Araştırmanın sonuçları, öngörülerine uygun şekilde, fazla yoksulluğun evdeki fazla düzensizlikle bağlantılı olduğunu gösterdi (b yolu). Ayrıca fazla düzensizlik, fazla psikolojik sıkıntıyla bağlantılıydı (c yolu). Son aşamada, b ve c yollarını hesaba katan bir istatistik işlemi olan iz analizi, başlangıçta bulunan a yolu (yoksulluk ve sıkıntı arasındaki) ilişkisinin aslında sıfır (yani ilişki yok) olduğunu gösterir. Evans ve arkadaşlarının bulduğu şey de tam olarak buydu. İz analizi onlara, yoksulluk ile çocukların psikolojik sıkıntıları arasındaki ilişkinin evdeki kaosun derecesi ile (ya da yönlendiriciliğiyle) açıklanabileceğini söyleyebilme olanağı sağladı.

Evans ve arkadaşları olası yönlendirici değişkenleri açıklamamışlardır ama biz hipotetik bir örnek verebiliriz. Şekil 5.7'deki ilişki örüntülerinin kız çocuklarıyla karşılaştırıldığında erkek çocukları için farklı olduğunu varsayalım. Örneğin, kaosun yönlendirici etkisinin sadece erkek çocukları için geçerli olduğunu, kız çocukları için olmadığını varsayalım. Bu durumda çocuğun cinsiyetinin, psikolojik sıkıntı, yoksulluk ve düzensizlik arasındaki ilişkilerin gücü ve yönünü etkileyen aracı bir değişken olduğunu düşünecektik. Diğer olası aracı değişkenlere nüfus yoğunluğu (örneğin şehir ya da köy) ve çocuğun kişiliğinin direnç derecesi (örneğin yüksek ya da düşük direnç) de eklenebilir. Yoksulluk, kaos ve psikolojik sıkıntı arasındaki ilişkilerin bu aracı değişkenlere göre nasıl değişeceğine dair hipotezler geliştirebilir misiniz?

İlişkisel araştırmalar, nedensel çıkarımlar için kesin bir dayanak sunmasalar da iz analizinde gözlenen ilişki örüntüleri, değişkenler arasındaki nedensel ilişkileri belirlemekte önemli ipuçları sağlar. Nedensel çıkarımlar yapmak isteyen araştırmacılar için bir sonraki adım, Bölüm 6 ve 7'de anlatıldığı gibi deneyler yapmaktır. Örneğin, kaosun (öngörülemeyen sonuçlar, gürültü gibi) laboratuvar ortamında değişimlenmesi, farklı ekonomik tabakalardan gelen bireyler arasında farklı sıkıntı düzeylerine neden olabilir. Bu çok yöntemli yaklaşım, psikolojik uyum ile yoksulluk arasındaki ilişkinin anlaşılmasında, kaosun nedensel rolüne ilişkin daha ayrıntılı bir kanıt bulmaya yardımcı olacaktır.

ÖZET

Anket araştırmaları, insanların özelliklerini (örneğin, demografik değişkenler) ve düşüncelerini, görüşlerini ve duygularını belirlemek amacıyla yeterli ve verimli araçlar sağlar. Ayrıca, doğal değişkenler arasındaki birlikte değişimler (korelasyon) değerlendirilerek yordayıcı ilişkiler belirlenebilir. Anketler, amaçları ve kapsamları bakımından farklıdır ama genellikle örnek seçmeyi gerektirir. Sonuçlar, dikkatlice seçilmiş örnek gruplardan sağlanır ve araştırılan bütün

evreni betimlemek için kullanılır. Anketlerde genellikle bir soru formu şeklinde önceden belirlenmiş bir dizi soru kullanılır.

Örnek seçme, evrenin tamamını temsil ettiği kabul edilen bir havuzdan, belirli sayıdaki birimin yansız olarak seçilmesi işlemidir. Örneklemden evrene genelleme yapma başarımız, örneklemin temsil ediciliğine (yani örneklemin evrenle aynı özellikleri göstermesine) bağlıdır. Temsil ediciliğe ulaşmanın en iyi yolu olasılığa dayanmayan örnekleme yerine olasılığa dayalı örnekleme kullanmaktır. Basit seçkisiz örneklemede, her birimin örnekleme girme olasılığı eşittir. Tabakalı seçkisiz örnekleme, evrenin alt kümeleri analiz edilmek istenildiği zaman kullanılır.

Dört genel anket yöntemi vardır: Posta anketleri, yüz yüze görüşmeler, telefon görüşmeleri ve internet anketleri. Posta anketleri görüşmeci yanlılığını önler ve özellikle mahrem veya rahatsız edici konuların araştırılması için uygundur. Tepki oranı yanlılığı nedeniyle oluşabilecek potansiyel sorunlar, posta anketlerine ciddi bir sınırlılık getirir. Yüz yüze görüşmeler ve telefon anketleri genellikle yüksek bir tepki oranı ve geniş bir esneklik sağlar. Telefon anketleri sıklıkla kısa araştırmalar için kullanılır. İnternet anketleri verimli ve düşük maliyetlidir; aynı zamanda, anket araştırmacılarına yeni fırsatlar sağlar. Bununla birlikte, örneklem yanlılığı ortaya çıkabilir ve cevaplama ortamının denetlenememesi nedeniyle hem yöntemsel ve hem de etik sorunları artmaktadır.

Anket araştırmaları, araştırma deseni adı verilen genel bir plana göre yapılır. Üç anket araştırma deseni vardır: Kesitsel desen, ardışık bağımsız örneklemler deseni ve boylamsal desen. Kesitsel desen, evrenin özelliklerini belirlemeye veya belirli bir anda iki veya daha fazla evren arasındaki farkları keşfetmeye odaklanır. Görüşler veya tutumlar arasında zamanla ortaya çıkan değişimleri belirlemek için, ardışık bağımsız örneklemler deseni veya boylamsal desen kullanılması gerekir. Boylamsal desenin tercih edilmesinin nedeni, araştırmacıların belirli bireylerde zamanla ortaya çıkan değişimi gözlemelerine ve denk olmayan ardışık örneklemlerin sorunlarından kurtulmaya olanak sağlar.

Anket araştırmalarının öncelikli veri toplama aracı, soru formlarıdır. Soru formları demografik değişkenleri ölçmek ve insanların tercih ve tutumlarını değerlendirmek için kullanılır. Kurallara uygun bir soru formu hazırlamak için araştırmacılar hangi bilgilere ihtiyaçları olduğuna, anketin nasıl uygulanacağına ve soruların hangi sırada daha etkili olacağına karar vermelidir; bu koşullar, anketin geçerli ve güvenilir olmasını sağlayacaktır. En önemlisi, sorular belirli, açık ve olabildiğince tek anlamı olacak şekilde yazılmış olmalıdır.

Diğer sözlü raporların sonuçları gibi anket sonuçları da, aksine bir kanıt olmadıkça (örneğin, sosyal açıdan kabul edilebilir cevaplar vermesi için katılımcıya baskı yapılmadıkça), olduğu gibi kabul edilebilir. İnsanların davranışları, yaptıklarını söyledikleri şeylerle daima uyumlu olmaz; bu nedenledir ki anket araştırmaları hiçbir zaman doğrudan gözlemin yerini tutmaz. Böyle olmakla birlikte, anket araştırmaları insanların tutumlarını ve düşüncelerini incelemek için mükemmel bir olanak sağlar.

İlişkisel (korelasyonel) bulguların yorumlanmasında en büyük sorun, nedensellik ve korelasyon arasındaki ilişkilerin anlaşılmasıdır. İki değişken arasındaki bir ilişki, bu iki değişken arasında nedensel bir bağlantı olduğunu göstermeye yetmez. Bununla birlikte ilişkisel kanıtlar, karmaşık istatistik

tekniklerle (iz analizindeki aracı ve yönlendirici analizler gibi) ve çok yöntemli yaklaşımlarla birlikte kullanıldığı zaman nedensel ilişkilerin belirlenmesine katkıda bulunabilir.

TEMEL KAVRAMLAR

korelasyonel araştırmalar 136
evren 139
örneklem 140
seçim yanlılığı 140
temsil edicilik 140
olasılığa dayalı örnekleme 141
olasılığa dayanmayan örnekleme 141
basit seçkisiz örnekleme 142
tabakalı seçkisiz örnekleme 143
tepki oranı yanlılığı 145
görüşmeci yanlılığı 146
kesitsel desen 151
ardışık bağımsız örneklem deseni 152
boylamsal desen 154
anket 158
sosyal beğenirlik 168
sahte ilişki 170

GÖZDEN GEÇİRME SORULARI

1 Anket araştırmasının amacını ve anket araştırmasında korelasyonun nasıl kullanıldığını kısaca tanımlayın.
2 Anketi yaptıran firmanın sonuçlarla ilgili beklentisinin, anket sonuçlarında yanlılığa neden olup olmadığını anlamak için hangi bilgilerinizi kullanırsınız?
3 Anketlerin, hazırlanış amaçlarından bağımsız olarak, iki ortak özelliği nedir?
4 İçinden bir örneklem grubunun seçileceği bir evrenin homojenliği (benzeşikliği) ile temsil özelliği olan bir örneklem büyüklüğü arasındaki ilişkiyi açıklayınız.
5 Bir kolaylık örneklemi (convenience sample) kullanılmasının, anket sonuçlarının yorumlanması için neden bir tehdit olduğunu açıklayınız.
6 Anket araştırması projeniz için posta uygulamasını, kişisel mülakatları, telefon mülakatlarını veya internet anketini niçin tercih edeceğinizi açıklayınız.
7 Sadece tepki oranının %50 olduğunu bilmenin, bir örneklemin bir evreni temsil etmediği çıkarımını yapmaya neden yetmediğini açıklayın..
8 İnternet anketlerinin başlıca avantajları ve dezavantajları nelerdir?
9 Bir ardışık bağımsız örneklemler deseninde, bir evrendeki tutum değişikliklerinin zamanla nasıl değiştiğini belirleyebilmek için, örneklemler arasında nasıl bir ilişki bulunmalıdır?
10 Bireylerin görüşlerinde zamanla meydana gelen değişimlerin büyüklüğünü ve yönünü değerlendirmek istiyorsunuz. Hangi anket yöntemini seçersiniz? Neden?
11 Bir öz bildirim ölçümünün güvenirliğini belirlemek için bir yöntem ve geçerliliğini belirlemek için bir yöntem belirtin.
12 Bir anket araştırmasında öz bildirim ölçümlerinin güvenirliğini etkileyebilecek üç faktörü açıklayın.
13 Birisi size anket sonuçlarınızın işe yaramaz olduğunu zira katılımcıların anketteki soruları dürüstçe yanıtlamadığını söylediğinde nasıl bir yanıt verirdiniz.
14 Neden "Korelasyon nedenselliğe işaret etmez"? açıklayın ve korelasyona dayalı kanıtların, davranışın muhtemel nedenlerini belirlemede nasıl faydalı olduğunu açıklayın.
15 *Aracı* ve *yönlendirici* değişkenleri tanımlayın ve her birine birer örnek verin.

DÜŞÜNDÜRÜCÜ SORULAR

1 Anket araştırmasını iyi bir şekilde yapabilmek zordur; özellikle de konu kişilerin cinsel tutum ve davranışları olduğunda bu böyledir. Kısmen kadın cinselliğine odaklanmış bir kitap için bir araştırmacı 43 eyaletteki çeşitli kadın gruplarına mensup kadınlara 100.000 anket postalamıştır. Bu gruplar, feminist organizasyonlardan kilise ve bahçıvanlık gruplarına kadar değişmektedir. Ankette 127 adet açık uçlu soru vardır. Araştırmacıya 4500 tane kadın dönüş yapmıştır.

Bulgular, 5 yıl veya daha fazla süre evli olan katılımcıların %70'inin evlilik dışı ilişki yaşadıklarını ve bu bunların %95'inin ise sevdikleri adamın duygusal tacizine uğradıklarını göstermiştir.

A Bu çalışmanın örneklemi büyüktür (4500). Bu büyüklük, örneklemin temsil gücü olduğunu garanti eder mi? Eğer değilse hangi potansiyel anket araştırması problemi bu örneklemin temsil kabiliyetini azaltmış olabilir?

B Bu sorunun A kısmına vermiş olduğunuz yanıta göre, araştırmacının yaptığı çıkarımların yanlış olduğunu iddia etmek mümkün müdür?

2 Elektronik haberleşmenin ve internetin artan kullanımıyla birlikte gençler arasında büyüyen bir sorun "internette taciz" deneyimidir (siber zorbalık). Siber zorbalığın yaygınlığına dair tahminler farklıdır. 2005 yılında yapılan ulusal bir taramada (ABD) profesyonel görüşmeciler, İngilizce konuşan ve evlerinde internet erişimi olan 10-17 yaşları arasında ergen bulunan 1500 ev sahibinden oluşan bir örneklemle rastgele telefonda görüşmüşlerdir (Gençlik İnternet Güvenliği Anketi). Ebeveynleri onay veren gençlerle de görüşülmüş ve gencin de onayı alınmıştır. Bu örneklemdeki gençlerin yüzde dokuzu önceki bir yıl içerisinde "online" olarak tacize uğradıklarını belirtmişlerdir. 2008 yılında yapılan diğer bir çalışmada ise Boston'da oturan 20.406 lise öğrencisi (9 ila 12. sınıflar) okulda bir gün sağlık ve davranış konusundaki bir anketi isimsiz olarak doldurmuştur. Sorulardan biri "İnternetten, telefondan veya başka bir elektronik haberleşme kanalından kaç defa taciz ya da tehdit edildiniz?" idi. Bu örneklemde öğrencilerin %15.8'i önceki bir yıl içerisinde mağdur edildiklerini bildirdiler.

Aşağıdaki ifadelerin her birisinin doğru mu yoksa yanlış mı olduğuna karar verin, sonra da yanıtınızı açıklayın.

A Boston anketindeki çok daha büyük bir örneklemin gösterdiği %15.8'lik oran, ABD genç nüfusunun siber zorbalığa maruz kalan kesimine dair daha güvenilir bir tahmindir.

B Bildirilen siber zorbalık miktarları arasındaki farklılığın nedeni, her iki örneklemdeki yaş gruplarının farklı olması (10-17 ve 9-12. sınıflardakiler) olabilir.

C Boston çalışmasındaki gençlerin daha fazla siber zorbalık bildirmesinin nedenlerinden biri, ulusal örneklem ile yapılan telefon görüşmesinden farklı olarak öz bildirim anketinin kullanılması olabilir.

D Her iki anket çalışmasındaki internet zorbalığına dair bulgular, Boston çevresinde ulusal ortalamadan daha fazla internet zorbalığı olduğunu göstermektedir.

E Her iki anketin de bulguları 2005-2008 arasında siber zorbalık oranının nerdeyse %7 oranında arttığını göstermektedir.

3 Büyük bir holdingin insan kaynakları bölümü, çalışanların alternatif çalışma saatleri ile ilgili bilgilerini (örneğin, esnek çalışma saatleri, annelik-babalık saatleri gibi), ayrıca, şirketin bu işbirliği politikalarının adilliği hakkında ne düşündüklerini anketle araştırmak istemiştir. Ankette çalışanlara ayrıca, alternatif bir çalışma tarifesi isteyip istemedikleri de sorulmuştur. Beş bin tam zamanlı çalışanın yer aldığı bir maaş bordrosu listesinden, tabakalandırma yöntemiyle 1000 kişilik bir örneklem seçilmiştir. Tabakalar, holdingde kaç yıl çalışıldığına göre belirlenmiş ve oranlı olarak örneklenmiştir. Anketler kurum içi elektronik posta yolu ile örneklem grubuna yollanmış ve isimsiz olarak bir kutuya atılarak geri dönüşü yapılmıştır. Soru listesi 600 çalışan tarafından doldurularak geri dönüş yapılmıştır. Sonuçlar, yanıtlayanlardan 200'ünün (%33) farklı çalışma programlarına dair kurum politikalarını adil bulmadığını göstermiştir.

A Başlangıçtaki 1000 kişilik örneklem 5000 kişilik tüm tam zamanlı çalışan kadrosunu temsil ediyor muydu? Neden ediyordu veya neden etmiyordu?

B Farz edin ki insan kaynakları bölümünün başındaki kişi, cevap oranı yüzünden ve örneklemin üçte birinin kurum politikalarını adaletsiz bulması nedeniyle örneklemin yanlı olduğuna karar vermiştir. Bu doğru bir karar mıdır? Neden doğrudur veya neden doğru değildir?

C Farz edin ki anketten aşağıdaki sonuçlar elde edilmiştir:

		Hiç alternatif bir çalışma tarifesi talep ettiniz mi?	
		Evet (% 42)	*Hayır (% 58)*
Tarife politikalarına dair inanış:	*Adaletsiz*	150	50
	Adaletli	100	300

Ayrıca, insan kaynaklarından gelen veriler, genel anlamda, 1250 tam zamanlı çalışanın (toplam tam-zamanlı çalışanların %25'i) alternatif çalışma programları talep ettiğini işaret etmektedir. Bu verilere dayanarak örneklemin yanlı olup olmadığına dair nasıl bir sonuca varırsınız?

4 Farz edin ki bir hastanede hastaların hastanede yatarken gördükleri bakıma yönelik tutumları ile ilgili bir anket geliştirmekle görevlendirilmiş bir stajyersiniz. Hastane yönetimi geçen ay boyunca yataklı tedavi görerek taburcu edilmiş ve takip anketleri için iletişime geçilmeyi kabul etmiş hastaların bir listesini temin ediyor. Yöneticiler, hastaların hastanedeki süreçler (örneğin, yatış/taburcu, ilaçlar, fatura) ile ilgili memnuniyet düzeylerinin taburcu edildikten sonraki 2, 6 ve 12'nci aylar boyunca değişip değişmediğini bilmek istiyorlar.

A Yöneticilerin sorularını yanıtlamak için kullanılabilecek iki anket çalışması tasarlayın ve açıklayın. Bu tasarılardan her birinin bu proje kapsamında nasıl uygulanacağını ve her birinin avantajlarını ve olası kısıtlılıklarını belirtin.

B Hastane çalışanları anket yöntemi için üç yaklaşım öneriyor: (1) Hasta listesinden seçkisiz bir örneklem seçerek soru listesini uygulamak için telefon anketi yapın; (2) Seçkisiz seçilmiş bir hasta örneklem grubuna, içinde hastanenin internet sitesindeki soru listesinin linki bulunan bir mektup gönderin; (3) Hastanenin internet sitesindeki anketle ilgili bir duyuru postalayın ve hastalardan soruları internet sayfasında yanıtlamalarını isteyin. Her bir yaklaşımın avantajları ve kısıtlılıklarını belirleyerek en iyisi olduğunu düşündüğünüz yaklaşıma dair önerinizi ve gerekçenizi açıklayın.

C Hastalara verilecek soru listesinde kullanılabilecek bir serbest (açık uçlu) ve bir kapalı uçlu soru örneği verin. Bu örnekleri her bir soru tipinin avantajlarını ve dezavantajlarını açıklamak için kullanın.

İleri Alıştırma I'in Cevapları

Seçim yanlılığı bütün bu hipotetik örneklerde geçerlidir. Bunu görmenin en iyi yolu, örneklem seçmede kullanılan işlemlerin, yürütülen araştırmada ulaşılan sonuçlara izin verip vermediğini sorgulamaktır.

1 O bölgedeki psikologların katılabilecekleri bölgesel toplantılar yapılabilir ve psikoloji alanı içindeki alt bölümlere odaklanılabilir (örneğin, araştırma psikologları, klinik psikologlar, adlî psikologlar gibi). Araştırmacının örnek seçme yöntemine yanlış diyemeyiz fakat sadece bir psikologlar grubundan gelen psikologları seçtiği düşünülürse, vardığı sonuçtan emin olunamaz. Düzenli olarak birliklere katılım göstermeyen psikologlar örneklemin içinde yer almamıştır. Sonuçlar, belki bölgesel birliklere katılım gösteren psikologları temsil edebilir ama tüm profesyonel psikologlara genellenemez.

2 Yeni başlayan spor psikoloğu, bu özel oyuna gelen kişilerin isim değişikliği konusunda neler düşündüğü hakkında bir çıkarım yapabilir. Fakat kullandığı işlem, çıkardığı sonuçları bölgesel spor fanatiklerine genellemesine izin vermez. Bölgesel spor fanatikleri farklı takımların fanatiklerini ve hatta oyunlara hiç gitmeyen kişileri bile kapsar.

3 Bu gazete muhabiri resmî oylama sonuçlarının nasıl ortaya çıkacağını bulmuş olabilir. Fakat sonuçları "kampüsteki öğrencilerin" görüşlerini mutlaka yansıtmak zorunda değildir. Bu oylamaların büyük bir çoğunluğu, sunulan öneriyle aynı görüşte olmuş olabilir. Fakat ne kadar öğrencinin oy kullandığını veya oy kullanmayan öğrencilerin düşüncelerini bilmiyoruz. Bazı öğrenciler birilerine ne yönde oy kullandıklarını söylemeyi reddetmiş olabilirler. Aynı problem, büyük politik seçimlerde ilk sonuçlar alındığı zaman ve sandık görevlileri kendi sonuçlarının tüm nüfusu temsil ettiğini belirttikleri zaman da ortaya çıkar.

4 Sinemaya gelenler ekstra kredi ile ödüllendirilme yöntemini tercih eden öğrenciler olabilir. Sinemaya gelmeyen öğrencilerin tercihlerini bilemeyiz.

İleri Alıştırma II'in Cevapları

1 Birinci öğrenci araştırmacı 100 "Yunan" ve 100 "bağımsız" öğrencinin içinden tabakalı seçkisiz örnekleme ile seçmeyi önermişti. Bu planın çok ciddi potansiyel bir kusuru, örneklemin evren içindeki Yunanlılar ile bağımsızların oranlarını temsil edememe olasılığıdır (sırasıyla %25 ve %75). Bu durum örneklem yanlılığına neden olur çünkü Yunanlılar araştırmada fazla temsil edilmiş olur. İkinci öğrenci araştırmacı, kampüs evreninin içinde 100 öğrenciden oluşan basit seçkisiz örneklem almayı önermiştir. Bu daha temsil edici bir örneklem olasılığını sağlasa bile Yunanlı kategorisinde onların görüşünü temsil edecek katılımcıların sayısını düşürebilir (yaklaşık 25 Yunanlı beklenir).

2 Tercih edilen örneklem planı, tabakalandırılmış seçkisiz örnekleme yöntemi kullanıyor. Bu örneklemede Yunanlılar ve bağımsızlar için örneklem büyüklüğü, oran olarak evrendeki değerlere uygundur. Örneklemdeki 200 öğrencinin 150'sini bağımsız örneklem havuzundan, 50'sini ise Yunanlı örneklem havuzundan seçersiniz.

Düşündürücü Soruların Cevapları

A Genelde daha büyük örneklem, evreni daha iyi temsil eder. Bu araştırmadaki problem, son örneklemin (geniş olmasına rağmen) 100.000'lik orijinal örnekleme göre düşük yanıt oranı vermesidir (%4.5) Düşük yanıt oranı ve araştırmanın konusu, anketi sadece çok motive olan kadınların tamamlamış olması olasılığını ortaya çıkarır ve 4500 kadından oluşan bir örneklemin tüm kadın evrenini temsil etmesi olasılığını düşürür.

B Cevap oranı çok düşük olduğu için, varılan sonuçların doğru ya da yanlış olduğunu söylemeye izin vermez. Bu anketin sonuçlarının doğru olduğunu belirlemenin, uygun en az bir yolu vardır. Literatürdeki kadınların cinsel tutum ve pratikleri üzerine yapılmış 1 veya daha fazla araştırmanın sonuçlarını toplamanız gerekir. Bu diğer araştırmaların kullanmış olduğu kadın örneklemlerini bilmek çok önemlidir. Ancak ondan sonra bu araştırmanın sonuçlarını, diğer araştırmanın sonuçlarıyla karşılaştırabilirsiniz. Eğer temsil edici örneklemleri olan araştırmaların sonuçları ile mevcut araştırmanın sonuçları örtüşüyorsa, mevcut araştırmanın sonuçlarının doğru olduğunu söyleyebiliriz. Elbette, kendiniz de bu araştırmada var olan problemleri ortadan kaldırdığınız bir araştırmayı yürütebilir ve kendi sonuçlarınızın bu araştırmadakilerle uyumlu olup olmadığını belirleyebilirsiniz.

KISIM ÜÇ

Deneysel Yöntemler

İSTANBUL TİCARET ÜNİVERSİTESİ
KÜTÜPHANESİ

BÖLÜM ALTI

Bağımsız Grup Desenleri

BÖLÜMÜN ANA HATLARI

GENEL BAKIŞ
PSİKOLOGLAR NEDEN DENEY YAPAR
DENEYSEL ARAŞTIRMANIN MANTIĞI
SEÇKİSİZ (YANSIZ) GRUPLAR DESENİ
Bir Seçkisiz Gruplar Deseni Örneği
Blok Seçkisizleştirme
İç Geçerliğe Tehditler
DENEYSEL BULGULARIN ANALİZİ VE YORUMLANMASI
Deneylerde Veri Analizinin Rolü
Sonuçların Betimlenmesi
Verilerin Anlamının Onaylanması
Veri Analizi Bize Ne Söyleyemez
DENEYSEL BULGULARIN DIŞ GEÇERLİĞİNİ SAĞLAMA
EŞLEŞTİRİLMİŞ GRUPLAR DESENİ
DOĞAL GRUPLAR DESENİ
ÖZET

GENEL BAKIŞ

Bölüm 2'de psikolojide araştırmanın dört amacından bahsettik: Betimleme, kestirim (prediction), açıklama ve uygulama. Psikologlar davranışı betimlemek için sıklıkla doğal ortamlarda gözlem yöntemlerini kullanırlar. Anketler psikologların, insanların tutum ve düşüncelerini betimlemelerine olanak sağlar. Psikologlar birlikte değiştiğini keşfettikleri ölçümler ve gözlemler yaptıkları durumlarda, davranış ve bilişsel süreçler hakkında kestirimler yapabilirler. Betimleme ve kestirim, davranışı bilimsel olarak betimlemek için zorunludur ama davranışın nedenlerini anlamak için yeterli değildir, çünkü psikologlar davranışın "nedenini" de açıklamak isterler. Bir olayın nedenlerini belirlediğimiz zaman, bilimsel açıklamaya ulaşmış oluruz. Bölüm 6, 7 ve 8 nedensel ilişkileri tanımlamak için en uygun araştırma yöntemine odaklanmaktadır *–deneysel yöntem.* Bu bölümlerde pratik önemi olan soruları cevaplamak kadar psikolojik teorileri test etmek için de deneysel yöntemin nasıl kullanıldığını araştıracağız.

Daha önce de belirttiğimiz gibi, araştırmada yaygın kullanılan yaklaşım *çok yöntemli* yaklaşımdır. Farklı yöntemler kullanarak bir araştırma sorusuna birbiriyle karşılaştırılabilir sonuçlar elde ettiğimiz zaman, bu sonuçlarımıza daha çok güvenebiliriz. Bu durumda sonuçlarımızın *"ortak bir noktada birleşen geçerliğinin (convergent validity)"* olduğunu söyleyebiliriz. Her yöntemin zayıf yönleri farklıdır ama bu kusurları telafi edebilecek güçlü yanları da vardır. Deneysel yöntemin özel gücü, neden -ve- sonuç ilişkisi kurabilmesidir. Bu bölümde bilim insanlarının neden deney yaptıklarını ve deneysel araştırmanın mantığını tartışacağız. Yaygın şekilde kullanılan bir deneysel desene odaklanacağız –seçkisiz gruplar deseni. Seçkisiz grupları oluşturmak için kullanılan işlemleri ve özellikle seçkisiz gruplar deseninin yorumuna yönelik tehditleri açıklayacağız. Daha sonra araştırmacıların, elde ettikleri sonuçları analiz etmek ve yorumlamak için kullandıkları işlemleri açıklayacağız; ayrıca, deneysel bulgularının dış geçerliğini nasıl belirlediklerini ele alacağız. Bu bölümde son olarak, iki ek bağımsız gruplar desenini inceleyeceğiz eşleştirilmiş gruplar deseni ve doğal gruplar deseni.

PSİKOLOGLAR NEDEN DENEY YAPAR

- Araştırmacılar, davranışın nedenleri hakkındaki hipotezlerini test etmek için deneyler yaparlar.
- Deneyler araştırmacıların, bir uygulamanın veya programın davranışı etkili bir şekilde değiştirip değiştirmeyeceğine karar vermesine olanak sağlar.

Psikologların deney yapmalarının temel nedeni, psikolojik teorilerden elde ettikleri hipotezleri görgül (ampiric) olarak test etmektir. Örneğin, önde gelen bir duygu ve stres teorisine göre, acı veren deneyimlere ilişkin düşünce ve duyguları tutmanın fiziksel bir bedeli olur (Pennebaker, 1989). Bu "ketleme teorisine (inhibition theory)" göre, kişinin bu deneyimleri kendinde tutması fiziksel stres yaratır.

Bu teoriyi test eden bir çok deneyde, bir gruptan kişisel duygusal olayların yazmaları istenirken, diğer gruptan yüzeysel konular hakkında yazmaları istendi. Sonuçlar, ketleme teorisiyle uyumlu çıktı: Duygusal konularda yazan katılımcıların sağlık sonuçları, yüzeysel konularda yazan katılımcıların sonuçlarından daha iyi çıktı. Ama yine de tüm sonuçlar ketleme teorisi ile uyumlu değildi. Örneğin, kendilerinden duygusal yaşantılarını dışa vurarak dansetmesi istenen öğrenciler, yaşantılarıyla ilgili olarak hem dans eden hem de yazan öğrencilerin elde ettikleri sağlıkla ilgili faydaları sağlayamadılar. Böylece, dışavurum pozitif sonuçlar için yeterli olamadı. Teori daha ileri düzeyde test edildiği zaman bulundu ki duygusal yaşantıların yazılması sırasında ortaya çıkan bilişsel değişiklikler, olumlu sağlık sonuçlarını açıklamada çok önemlidir (Pennebaker & Francis, 1996).

Ketleme teorisini test etmeye ilişkin kısa betimlememiz, psikologların bir teoriden elde edilmiş bir hipotezi test etmek için deney yaptıkları zaman izledikleri genel süreci gösteriyor. Eğer deneyin sonuçları, hipotezin beklentisi ile uyumlu ise, o zaman teori desteklenmiş olur. Ama sonuçlar beklenenden farklı ise, o zaman teori üzerinde değişiklikler yapılması ve teorinin yeni şekline uygun yeni bir hipotez geliştirilmesi ve bir başka deneyde test edilmesi gerekebilir. Deneylerin sonuçlarına dayanarak hipotezleri test etmek ve teorileri yeniden düzenlemek, bazen uzun ve zahmetli bir süreç olabilir; tıpkı bir resmi tamamlamak için bir bulmacanın parçalarını bir araya getirmek gibi. Deneyler ile öne sürülen açıklama arasındaki etkileşim, psikologların düşünme, hissetme ve davranış biçimimizin nedenlerini anlamak için kullandıkları temel bir araçtır.

İyi yürütülen deneyler çok çeşitli alanlarda, tedavilerin etkisi hakkında önemli bilgiler sağlayarak toplumun sorunlarını çözmeye yardım eder. Deneylerin bu rolü tıp alanında uzun bir geçmişe sahiptir (Thomas, 1992). Örneğin, 19'uncu yüzyıl başlarında tifo ateşi ve titremeli hezeyan (delirium tremens) çoğunlukla öldürücüydü. Bu dönemde bu iki durumu tedavi etmek için standart olarak kanatma, temizleme gibi teknikler uygulanırdı. Bir deneyde bu uygulamaların etkinliğini test etmek için bir gruba standart uygulama (kanatma, arındırma) yapıldı, diğerine bir işlem yapılmadı, sadece istirahat, iyi beslenme ve yakın gözlem uygulandı. Thomas (1992) yaptığı deneyin sonuçlarını "net ve ürkütücü" olarak tanımladı (s. 9): Standart tıbbi tedavi alan grup tedavi almayan gruba kıyasla daha kötü bir performans sergiledi. Yani, 19'uncu yüzyıl tekniklerini kullanarak bu gibi durumları tedavi etmeye çalışmak, onları hiç tedavi etmemekten daha kötü sonuçlara yol açmaktaydı. Bu tür deneyler bir çok tıbbî koşulun sınırlılığına dair iç görüye katkı sağladı: Hastalık doğal seyrini izler ve hastalar kendi kendilerine iyileşirler.

DENEYSEL ARAŞTIRMANIN MANTIĞI

- Araştırmacılar bir deneyde bağımsız değişkeni, davranış üzerindeki etkisini gözlemek için değişimlerler; bu etkinin varlığı, bağımlı değişkenle ölçülür.
- Deneysel kontrol araştırmacıların, bağımlı değişkende gözlenen değişikliklere, bağımsız değişkenin *neden* olduğuna ilişkin nedensel çıkarım yapmalarına olanak sağlar.

- Kontrol, deneylerin vazgeçilmez unsurudur; deneysel kontrol değişimleme, koşulların sabitlenmesi ve dengeleme yolu ile gerçekleştirilir.
- Bir deneyde nedensel çıkarım için şu üç koşulun sağlanması gerekir: Birlikte değişim (covariation), birbirini izleme (time order) ve olası alternatif açıklamaların elenmesi (elimination); bu üç koşul gerçekleştiği zaman iç geçerliğin de sağlandığı sonucuna varılır.
- Bozucu (confounding) etkiler var ise, gözlenen birlikte değişimi açıklayabilecek alternatif açıklamaların olduğu düşünülür; bu durumda deney, iç geçerliğini kaybeder. Olası alternatif açıklamalar, koşulları sabit tutarak ve dengeleyerek ortadan kaldırılır.

Gerçek bir deney bir ya da bir kaç faktörü *değişimlemeyi* ve bu değişimlemenin davranış üzerindeki etkilerini *ölçmeyi* (gözlemini) kapsar. İkinci Bölüm'de görüldüğü gibi, araştırmacılar tarafından kontrol edilen ya da değişimlenen faktörlere *bağımsız değişken* denilir. Bir bağımsız değişkenin en az iki düzeyi (koşul) olmalıdır. Bir düzey "uygulama" koşulu, diğer düzey kontrol (ya da kıyaslama) koşuludur. Sıklıkla, gruplar arasında ek karşılaştırmalar yapmak için ikiden fazla düzey kullanılır. Eğer varsa, bağımsız değişken(ler)in etkilerini gözlemek için kullanılan ölçümlere *bağımlı değişken* denilir. Bu iki değişken türü arasındaki ayrımı hatırlamanın bir yolu da bağımlı değişkenin bağımsız değişkene *bağlı olduğunu* anlamaktır.

Deneyler hipotezlerin test edilmesinde etkilidir, çünkü bizim belirli bir koşulda oldukça yüksek bir kontrol sağlayarak çalışmamızı mümkün kılar. Araştırmacılar bağımsız değişkenin bağımlı değişken üzerinde gözlenebilir etkilere *neden olduğundan* emin olmak için, deneyleri kontrollü bir biçimde yaparlar. Nedensel çıkarımı yapabilmek için gerekli üç koşul; birlikte değişim, zamansal ardışıklık ve alternatif açıklamaların elenmesidir (bk. Bölüm 2).

Bir deneyin bağımlı ve bağımsız değişkenleri arasında bir ilişki gözlemlediğimiz zaman, birlikte değişim (covariation) koşulu karşılanmış olur. Araştırmacılar bir bağımsız değişkeni değişimler ve *sonra* davranışta bir değişim (değişimlemeye bağlı bir değişim) gözlemlenirse, zaman-sıra koşulu karşılanmış olur. Nihayet, kontrol koşullarının sağlanmasıyla, en önemlisi, *koşulların sabit tutulması ve dengelemeyle*, alternatif açıklamaların elenmesi koşulu yerine getirilmiş olur. Nedensel bir çıkarım için bu üç koşul karşılandığı zaman deneyin **iç geçerliğinin** olduğu söylenir; bu durumda, bağımsız değişkenin, bağımlı değişkende yaptığımız ölçüme göre, davranışta gözlemlediğimiz değişikliğe neden olduğunu söyleyebiliriz.

Temel Kavram

SEÇKİSİZ (YANSIZ) GRUPLAR DESENİ

- Bağımsız gruplar deseninde, her denek grubu bağımsız değişkenin sadece bir düzeyinde yer alır
- Koşullara seçkisiz (random) atama, denk gruplar oluşturmak için kullanılır; seçkisiz atama bunu, katılımcıların özelliklerini (bireysel farkları) bağımsız değişkenin farklı koşulları arasında dengeleyerek veya ortalamasını alarak yapar.
- Bir deneyde bağımsız değişkenin farklı düzeylerinin uygulanacağı bağımsız grupları oluşturmak için seçkisiz atama kullanıldığı zaman, bu deneye seçkisiz gruplar deseni denir.

Temel Kavramlar

Bir bağımsız gruplar deseninde, her denek grubu bağımsız değişkenin farklı bir koşuluna katılır.[1] En etkin bağımsız gruplar deseni, bağımsız değişken uygulanmadan önce, denk gruplar oluşturmak üzere deneklerin koşullara **seçkisiz atandığı** desendir. Koşullara seçkisiz atama yapılıyorsa, bağımsız gruplar deseni **seçkisiz gruplar deseni adını** alır. Desenin mantığı çok açıktır. Deneyin başında, gruplar tüm önemli özellikleri açısından benzer olacak şekilde oluşturulur. Daha sonra, bağımsız değişkenin düzeyleri hariç, tüm gruplara aynı uygulama yapılır. Böylece, gruplar arasında bağımlı değişken açısından gözlenen herhangi bir farkın nedeninin, bağımsız değişken olması gerektiği sonucuna varılır.

Bir Seçkisiz Gruplar Deseni Örneği

Deneysel yöntemin mantığı ve iç geçerliği sağlayan kontrol tekniklerinin uygulanması, İngiltere'de Dittmar, Halliwell ve Ive'ın (2006) vücutlarından memnun olmayan kızlarla yaptığı bir deneyde gösterilmiştir. Onların amacı, çok zarif kadın fotoğraflarıyla karşılaşmanın, kızların kendi vücutları hakkında olumsuz duygular yaşamalarına neden olup olmadığını araştırmaktı. Ergenler ve yetişkin katılımcılarla yürütülen bir çok deneyin sonucuna göre kadınlar ince bir kadın modelle karşılaştıktan sonra, ince olmayan modellere kıyasla, kendi bedenlerinden daha fazla memnuniyetsizlik göstermiştir. Dittmar ve arkadaşları daha küçük yaşlarda aynı etkinin bulunup bulunmayacağını anlamak için, beş yaşındaki kız çocuklarına kadar indiler. Test edilen çok ince vücut imgesi Barbi bebekti. Barbi'nin vücut oranlarını gerçek kadınlarla karşılaştıran antropolojik çalışmalar, Barbi bebeğin gerçeği yansıtmayan vücut oranları olduğunu göstermiştir. Yine de Barbi kadın güzelliği için sosyokültürel bir ideal olmuştur (bk. Şekil 6.1).

Araştırmada küçük genç kız gruplarında (5½–6½ yaşında), kızlara "Mira" adında bir kızın alışverişe gittiği, elbiseler aldığı ve yaş günü partisine hazırlandığı bir hikâye okundu. Hikâyeyi dinledikten sonra kızlara resimli kitaptaki hikâye ile ilişkili altı resim gösterildi. Deneyin bir koşulunda, resimli kitapta hikâyenin farklı sahnelerinde Barbi görülmekteydi (örneğin, parti için alışverişte iken, partiye hazırlanırken). İkinci koşulda resimli kitapta benzer sahneler vardı, fakat resimdeki şekil "Emme" bebekti. Emme tipi bebek, ABD ölçülerine göre 16 beden olup, daha gerçekçi vücut ölçülerine sahipti (bk. Şekil 6.2). Nihayet üçüncü koşulda, resimli kitapta Barbi ya da Emme bebek (ya da herhangi bir vücut) resmi yoktu; bunun yerine, hikâye ile ilişkili nötr imgeler vardı (örneğin, elbise dükkanlarının pencereleri, renkli balonlar). Resimli kitaptaki bu üç tip model (Barbi, Emme, nötr) deneyde değişimlenen üç bağımsız değişken düzeyini temsil etmektedir. Bağımsız değişkenin her düzeyinde, farklı kız grupları yer aldığı için, bu deney deseni bağımsız gruplar deseni adını almaktadır.

[1]Bağımsız gruplar deseni için kullanılan bir diğer terim, gruplar arası desendir. Her iki terimde katılımcı grupların denk olduğu ama katılımcılar arasında herhangi bir örtüşmenin olmadığı gruplar için kullanılır (örneğin, her katılımcı sadece bir koşulda yer alır).

ŞEKİL 6.1 Amerika Birleşik Devletleri'nde 3-10 yaş arası kız çocuklarının en az bir tane Barbi bebeği vardır ve tipik bir kız çocuğunun sekiz tane Barbi bebeği vardır (Rogers, 1999).



Değişimleme (manipulation) Dittmar ve ark. (2006) kızların kendi vücutlarına ilişkin memnuniyetsizlikleri hakkındaki hipotezlerini test etmek için *değişimleme* kontrol tekniğini kullandılar. Bağımsız değişkenin üç düzeyi araştırmacılara hipotezlerine ilişkin karşılaştırmalar yapma olanağı verdi. Eğer sadece Barbi koşulunu test etselerdi, bu resimlerin kız çocuklarının kendi bedenlerine ilişkin memnuniyetsizliklerini etkileyip etkilemediğini belirlemeleri mümkün olmayacaktı. Böylece nötr resim koşulu, kızların beden memnuniyetsizliklerinin, baktıkları ince vb. nötr resimlere bağlı olarak farklılaşıp farklılaşmadığını belirleme olanağı sağladı. Emme koşulu önemli bir karşılaştırma olanağı sağladı. Herhangi bir beden imgesinin kızlarda kendilik algısını etkilemesi muhtemeldir. Bu araştırmacılar sadece Barbi ile temsil edilen *ince* vücut idealinin, beden memnuniyetsizliğine yol açacağı hipotezini test ettiler.

Hikâyenin sonunda, genç kızlar resimli kitaplarına döndüler ve kendi yaş düzeylerine göre düzenlenmiş bir anketi cevapladılar. Araştırmacılar kızlarda beden memnuniyetini değerlendiren birçok ölçüm kullanmış olsalar da, biz burada Çocuk Şekil Sıralama Ölçeği üzerinde duracağız.

ŞEKİL 6.2 Emme bebek 2002 yılında genç kızlar için gerçekçi bir beden oluşturmak üzere tanıtıldı. Bebek, Emme adında Amerika Birleşik Devletleri'ndeki bir süper modelin adını taşımaktadır.



Bu ölçek, çok inceden çok kiloluya doğru değişen bir aralıkta yer alan, iki sıra hâlindeki yedi kız bedeni çiziminden oluşmaktadır. Her kızın, ilk önce en üst sıradaki bedenlerden kendisinin şimdiki bedenine en çok benzeyen şekli boyaması istendi (algıladıkları kendi gerçek beden şekli). Sonra ikinci sırada, kızlara en fazla istedikleri görünümü temsil eden bedeni boyamaları istendi (kendilerinin ideal beden şekli). Kızlara bu şekillerden herhangi birini seçebilecekleri ve her iki sırada da aynı şekli seçebilecekleri söylendi. Her kızın gerçek şekli ve ideal şekli arasındaki şekil sayısı sayılarak, bir beden şekli memnuniyetsizlik puanı hesaplandı. Sıfır puan bedene ilişkin herhangi bir memnuniyetsizlik olmadığını, negatif puan daha ince olma isteğini ve pozitif bir puan daha kilolu olma isteğini göstermektedir.

Bu deneyin sonuçları açıktı: Barbi imgelerine maruz kalan genç kızlar, Emme imgelerine ya da nötr imgelere maruz kalan kızlara kıyasla kendi beden şekillerinden daha fazla memnuniyetsizlik yaşadılar. Emme imgesi ve nötr imgeye maruz kalınan koşulda, 20 kızın ortalama beden memnuniyetsizliği puanı sıfırdı. Tersine, Barbi imgesine maruz kalınan koşulda, 17 kızın ortalama memnuniyetsizlik puanı -.76 olup istekleri daha ince olmak yönündeydi. Bu deneyde

değişimleme yoluyla kontrol etme tekniği kullanılarak nedensel çıkarım için gerekli olan iki koşul karşılanmış oldu: (1) Kızların beden memnuniyetsizliğindeki farklar deneysel koşullarla birlikte değişti ve, (2) beden memnuniyetsizliği ideal ölçülerdeki bedenlerle karşılaştıktan *sonra* ortaya çıktı (zaman-sırası ilişkisi). Nedensel çıkarım için gerekli olan üçüncü koşul (alternatif açıklamaların elenmesi) ise, koşulların sabit tutulması ve dengeleme ile yerine getirildi.

Koşulların Sabitlenmesi Kızların kendi bedenlerine yönelik tutumlarını etkileyen bir kaç faktör, üç deneysel koşulda aynı tutuldu. Tüm kızlar alışveriş ve bir doğum günü partisi konulu bir hikâyeyi dinledi, ve resimli kitaplara aynı anda baktılar. Hepsi deney boyunca aynı yönergeleri ve deney sonunda aynı anketi aldı. Araştırmacılar, gruplar arasında ortaya çıkan sistematik farklılaşmanın nedeninin *sadece* bağımsız değişkenden kaynaklandığından emin olmak için *koşulları sabitleme* yöntemini kullanırlar.

Eğer üç grup katılımcının tutumu kitaptaki imgeler dışında bir faktörün etkisiyle değişmiş olsaydı, o zaman deney sonuçları açıklanamazdı. Varsayın ki Barbi deneysel koşulundaki katılımcılar farklı bir hikâye dinlediler; örneğin, Barbi'nin ince ve popüler olduğu hakkında bir hikâye. Kızlarda beden memnuniyetsizliğine ilişkin gözlenen farkın, Barbi imgelerinden mi yoksa bu farklı hikâyeden mi kaynaklandığını bilemezdik. Etkisi araştırılan bağımsız değişken ile farklı bir potansiyel bağımsız değişkenin birlikte değişimine izin verilirse, *bozucu* (confounding) etken ortaya çıkar. Karıştırıcı etkenler yok ise, bir deneyin *iç geçerliği* olduğunu söyleyebiliriz.

Koşulların sabitlenmesi, araştırmacıların karıştırıcı etkenleri önlemek üzere kullandıkları bir kontrol yöntemidir. Sabit tutulan bir etken, kızlara anlatılan hikâye gibi, deneyde değiştirilmez. Sabit olduğu için, ne bağımlı değişkenle ne de değişimlenen değişkenle birlikte değişemez. Böylece araştırmacılar bu sabit tutulan etkenlerin, gözlenen sonuçların potansiyel nedeni olmadığı sonucuna varırlar.

Yine de kontrol etmek istediğimiz, araştırılan davranışı etkilediğini düşündüğümüz *olası* alternatif nedenlerin farkında olmak önemlidir. Örneğin, kızların her koşulda duyduğu hikâye sabit tutulmuştur ama araştırmacılar muhtemelen oda sıcaklığı gibi durumları sabitlememiştir; çünkü oda sıcaklığının beden imgesini etkilemesi muhtemel değildir (en azından sadece bir iki derece değiştiğinde). Yine de deneylerimizde etkileri önceden kestirilemeyen ya da düşünülmemiş karıştırıcı etkenler olabilir.

Dengeleme Açıktır ki deneysel yöntemin bir başka temel noktası da deneyin başlangıcında karşılaştırılabilir (benzer) gruplar oluşturmaktır. Her gruptaki katılımcılar kişilik, zekâ ve benzeri özellikler açısından (bunlar aynı zamanda *bireysel farklar* olarak da bilinmektedir) benzer olmalıdır. Bu özelliklerin *dengeleme (balancing)* yöntemi ile kontrol edilmesi gerekir, çünkü sabitlenemezler. Seçkisiz atamanın amacı, deneysel koşullar açısından bireysel farkları

dengeleyerek birbirine denk gruplar oluşturmaktır. Dittmar ve arkadaşlarının (2006) kullandığı aşağıdaki seçkisiz gruplar deseni şöyle açıklanabilir:

Aşama 1	Aşama 2	Aşama 3
R_1	X_1	O_1
R_2	X_2	O_1
R_3	X_3	O_1

Tablodaki R1, R2, ve R3 harfleri deneklerin üç bağımsız deneysel koşula seçkisiz (Random) atandığını gösterir. X1, X2, ve X3 ise bağımsız değişkenin farklı düzeylerini temsil etmektedir; örneğin, X1 Barbi, X2 Emme ve X3 nötr resimleri temsil etmektedir. O1, O2, ve O3 ise, üç koşul sonundaki gözlemleri (Observation) temsil etmektedir.

Bu deneyde eğer Barbi imgesi ile karşılaşan katılımcıların, Emme ya da nötr imgelerin sunulduğu katılımcılara kıyasla, daha kilolu ya da daha çok Barbi bebekleri olduğu gösterilmiş olsa, bir başka alternatif açıklama ortaya çıkar: Barbi koşulunda daha fazla beden memnuniyetsizliği yaşamanın nedeni, daha kilolu olmak veya daha fazla Barbi bebek sahibi olmak olabilir (araştırmacının ifadesiyle bu, karıştırıcı bir etken olabilir). Benzer şekilde, kızların deney *öncesinde* var olan beden imgesine ilişkin bireysel farklar, çalışma bulgularını açıklamada akılcı alternatif bir açıklama olabilir. Böyle olmakla birlikte, gruplar arasındaki bireysel farkları dengelemek için seçkisiz atama kullanıldığı zaman, mantıksal olarak, bağımlı değişkende gözlenen bir farkın, katılımcıların özelliklerinden kaynaklanmış olması olasılığını eleyebiliriz.

Vücut ağırlığı gibi bir etkeni dengelersek, *ortalama* vücut ağırlığı cinsinden üç denk grup oluştururuz. Dikkat edilmelidir ki bu, tüm kızların vücut ağırlığının aynı olması demek olan vücut ağırlığının sabit tutulmasından farklıdır. Benzer şekilde, bu üç gruptaki kızların sahip olduğu Barbi bebek sayısını dengelemenin anlamı şudur ki üç grubun sahip olduğu bebeklerin *ortalama* sayısı aynıdır, her bir kızın sahip olduğu bebeklerin sayısı değil. Seçkisiz atamanın iyi yanı, sadece bizim belirttiklerimizin değil *tüm* bireysel farkların dengelenmesidir. Böylece katılımcılar arasındaki *herhangi bir* bireysel farka bağlı alternatif açıklamaları eleyebiliriz

Özetle, Dittmar ve arkadaşları ince beden imgelerine maruz kalmanın, genç kızların kendi bedenlerinden memnuniyetsiz olmalarına neden olduğu sonucuna vardı. Aşağıdaki sonuçlara varabildiler, çünkü;

- Kızların gördüğü imgelerin değiştirildiği bir bağımsız değişkeni değişimlediler.
- Bağımlı değişkeni etkileme olasılığı olan diğer etkenleri sabit tutarak, alternatif açıklamaları elediler.
- Seçkisiz atama ile gruplar arasındaki bireysel farkları dengelediler.

Kutu 6.1, seçkisiz gruplar deseninin genç kızların beden imgesini çalışmak üzere nasıl kullanıldığını özetlemektedir.

KUTU 6.1

KIZLARIN BEDEN İMGESİ DENEYİNİN ÖZETİ

Deneysel işlemin ana hatları: Yaşları 5½–6½ arasında değişen kız çocukları üç farklı gruba atandılar ve bir hikâye dinlerken, üç farklı resim kitabından birisine baktılar. Kitapları inceledikten sonra kızlar, kendi beden anlayışları konusunda sorulan soruları cevapladılar.

Bağımsız değişken: Katılımcıların baktığı resimli kitap türü (Barbi, Emme veya nötr imgeler)
Bağımlı değişken: Kızların gerçek beden imgeleri ile ideal beden imgeleri arasındaki fark ölçülerek belirlenen beden memnuniyetsizliği.
Kontrol İşlemlerinin Açıklaması
Koşulları sabit tutmak: Üç ayrı koşuldaki kızlara aynı yönergeler verildi, aynı hikâye dinlettirildi ve sonunda aynı sorular soruldu.
Dengeleme: Katılımcılar üç koşula seçkisiz atanarak, bireysel farklar dengelendi.
Nedensellik İçin Kanıt Sağlayan Deneysel Mantığın Açıklaması
Birlikte değişim: Kızların beden memnuniyetsizliğinin, deneysel koşulla birlikte değiştiği bulundu.
Zaman–sıra ilişkisi: Resimli kitap türü beden memnuniyetsizliğini ölçmeden önce değişimlendi.
Olası alternatif nedenlerin ortadan kaldırılması: Koşulları sabit tutarak ve seçkisiz atama yaparak bireysel farkları dengeleme, bozucu etkenleri eleyen bir kontrol sağladı.
Sonuç: Barbi resimleri gibi çok ince bir beden imgesine maruz kalmak, kızların kendi bedenlerinden memnuniyetsizlik duymalarına neden oldu.

[Dittmar, Halliwell ve Ive'den (2006) alınmıştır.]

İLERİ ALIŞTIRMA I

Bu alıştırmada, bir deneyin kısa bir açıklamasından sonra sorulan soruları cevaplamanız gerekiyor.

Bushman (2005) insanların reklam anılarının, izledikleri televizyon programının türüne göre değişip değişmediğini araştırdı. Katılımcılar (N = 336; yaş, 18–54 arası) şu dört tür televizyon programından birini izlemeleri için seçkisiz olarak atandılar: Şiddet içerikli (örneğin, Polisler), erotik (örneğin, Seks ve Şehir), şiddet ve seks içerikli (örneğin, CSI Miami), ya da nötr programlar (örneğin, Amerika'nın en ilginç hayvanları). Her TV programının içine 30'ar saniyelik 12 reklam yerleştirildi. Katılımcıların reklamlarda sunulan markalara eşit oranda maruz kaldıklarından emin olmak için, araştırmacılar yaygın olmayan markaları seçtiler (örneğin, "Dermoplast" "José Olé"). Her programın yaklaşık 12, 24 ve 36'ıncı dakikalarında dörder reklam verildi. Katılımcılar küçük gruplar hâlinde test edildiler ve her oturum katılımcıların rahat oturmalarını sağlamak üzere düzenlenmiş sandalyelerde yürütüldü; bu esnada hafif içecekler ve atıştırmalıklar sunuldu. Programı izledikten sonra, reklamlarla ilgili sürpriz bir bellek testi aldılar. Sonuçlar, seks ve şiddet içeren televizyon programlarında gösterilen reklamların daha az hatırlandığını gösterdi. En az hatırlanan reklamlar ise erotik içerikli programlarda gösterilenler oldu.

1 Bushman (2005) değişimleme yaparak, bu deneyin hangi yönlerini kontrol etti?
2 Bushman koşulları sabit tutarak bu deneyi hangi yönlerden kontrol etti?
3 Dengeleme yaparak deneyi hangi açıdan kontrol etti?

Bushman, B. J. (2005). Violence and sex in television programs do not sell products in advertisements. *Psychological Science*, 16, 702–708.

Blok Seçkisizleştirme (Block Randomization)

- Blok Seçkisizleştirme, denek özelliklerini ve yürütülen deney sırasında ortaya çıkan potansiyel karıştırıcıları dengeler ve denk büyüklükte gruplar oluşturur.

Temel Kavram

Seçkisiz atamayı (random assignment) yapmanın en yaygın işlemi **blok seçkisizleştirme**dir. Öncelikle blok seçkisizleştirmenin nasıl yapıldığını açıklayalım ve sonra ne olduğunu görelim. Varsayın ki A, B, C, D ve E ile sembolize edilen beş deneysel koşulun olduğu bir deney yapıyoruz. Bir "blok" beş koşulun seçkisiz sıralamasından oluşur.

Bir koşullar bloğu	→	Seçkisiz koşul dizisi
A B C D E		C A E B D

Blok seçkisizleştirmede, denekleri her seferinde bir blok koşuluna katılacak şekilde atarız. Beş koşulun olduğu örneğimizde, her koşulda beş deneğin ilk bloğu bir denekle tamamlaması gerekmektedir. Beş koşullu örneğimizde, ilk bloktaki her bir koşulu tamamlamak için beş katılımcıya ihtiyaç vardır. Sonraki beş denek ikinci bir bloğu oluşturmak için beş koşulun her birine atanır ve bu şekilde devam eder. Eğer beş koşulun her birinde 10 denek olmasını istiyorsak, o zaman blok seçkisizleştirme yöntemi ile 10 blok oluşturulması gerekir. Her blok beş koşulun seçkisiz olarak düzenlenmesinden oluşur. Bu işlem ilk 11 katılımcı için aşağıda gösterilmiştir.

10 Blok	*Katılımcılar*		*Koşul*	
1) C A E B D	1) Cara	→	C	Birinci Blok
2) E C D A B	2) Andy	→	A	
3) D B E A C	3) Jacob	→	E	
4) B A C E D	4) Molly	→	B	
5) A C E D B	5) Emily	→	D	
6) A D E B C	6) Eric	→	E	İkinci Blok
7) B C A D E	7) Anna	→	C	
8) D C A E B	8) Laura	→	D	
9) E D B C A	9) Sarah	→	A	
10) C E B D A	10) Lisa	→	B	
	11) Tom	→	D	

ve bu şekilde 50 katılımcı için seçkisiz olarak düzenlenmiştir.

Denekleri seçkisiz olarak gruplara atamak için blok seçkisizleştirme kullanılmasının bir çok avantajı vardır. İlk olarak, blok seçkisizleştirme eşit büyüklükte gruplar oluşturur. Bu önemlidir, çünkü her gruptaki gözlem sayısı her grup için betimleyici istatistiklerin güvenirliğini etkiler ve bu, elde edilen ölçümlerle gruplar arasında güvenilir bir karşılaştırma yapmak için arzu edilir bir özelliktir. İkinci olarak, blok seçkisizleştirme zamanla ilişkili değişkenleri kontrol eder. Deneylerin tamamlanması uzun zaman aldığı için,

bazı katılımcılar bu süre içinde meydana gelen olaylar tarafından etkilenebilir. Blok seçkisizleştirmede her koşul, her bir blokta test edilir; böylece zamanla ilişkili bu değişkenler deney koşulları boyunca dengelenir. Eğer blok seçkisizleştirme yöntemi kullanılıyorsa, örneğin bir deneyin yürütüleceği bir üniversite kampüsünde travmatik bir olay yaşanırsa, olayı yaşayan katılımcıların sayısı her bir koşulda denk olacaktır. Bu nedenle, olayın katılımcıların performansı üzerindeki etkilerinin denk ya da ortalama civarında olacağını varsayabiliriz. Blok seçkisizleştirme aynı zamanda araştırmacıların değişmesi ya da hatta deneklerin seçildiği evrenlerdeki değişiklikler gibi zamanla ortaya çıkabilecek diğer değişiklikleri de dengelemeye çalışır. Örneğin, eğer blok seçkisizleştirme kullanılırsa, hem güz hem de bahar yarıyıllarından katılımcı alınarak çok iyi bir deney yapılabilir. Blok seçkisizleştirme yönteminin iyi yönü, katılımcıların herhangi bir özelliğini (zamanla ilişkili olanlar dâhil) deney koşulları boyunca dengelemesidir.

Blok seçkisizleştirme işlemi üzerine alıştırma yapmak isterseniz, bu bölümün sonundaki düşündürücü sorularda yer alan 1A 'yı cevaplayın. Soru 1A bölümün sonundadır.

İç Geçerliğe Tehditler

- Grupları bağımsız değişkenin farklı koşullarına seçkisiz olarak atama, katılımcılar arasında önceden mevcut farklılıklar nedeniyle, potansiyel bir bozucu faktör yaratır.
- Blok seçkisizleştirme, dış (extraneous) değişkenleri bağımsız değişken koşulları arasında dengeleyerek, iç geçerliği arttırır.
- Mekanik denek kaybı değil ama seçici denek kaybı iç geçerliği tehdit eder.
- Plasebo kontrol grupları talep özelliklerini, çift-kör deneyleri ise hem talep özelliklerini hem de deneyci özelliklerini kontrol etmek için kullanılır.

Gördük ki *iç geçerlik*, başka bazı kontrol dışı bağımsız bir değişkendeki farklılıkların açık bir şekilde, kontrol dışı başka bazı değişkenlere değil, bağımsız bir değişkenin etkisine bağlayabilme derecesidir. Bu kontrol dışı değişkenler çoğunlukla iç geçerliği tehdit eder. Bu tehditler, bir araştırmanın bulgularının potansiyel alternatif açıklamalarıdır. Araştırmacılar bir bağımsız değişkene ait neden ve sonuç çıkarımı yapabilmek için **iç geçerliği tehdit eden etkenleri** kontrol ederler. Daha sonra, deneysel araştırmada iç geçerliğe tehdit oluşturan ve bu tehditleri kontrol etmeye yarayan çeşitli sorunları açıklayacağız.

Temel Kavram

Grupların Bütün Hâlinde (Intact Groups) Test Edilmesi Seçkisiz atama, seçkisiz gruplar deseninde denk gruplar oluşturmak için kullanılır. Yine de bazen, seçkisiz atama kullanılmış gözükse de, denk *olmayan* grupların oluştuğu durumlar vardır. Bu problem, gruplar *bütün hâlinde* (bireyler değil) yansız şekilde bir deneyin koşullarına atandığı zaman ortaya çıkar. Bütün hâlinde gruplar,

deney başlamadan önce oluşturulur. Örneğin, psikolojiye giriş dersi alan farklı sınıflar, bütün hâlinde gruplara örnektir; öğrenciler bu sınıflara seçkisiz olarak atanmazlar. Öğrenciler sıklıkla belirli bir sınıfta olmak isterler, bunun nedeni dersin zamanı, öğretmeni, sınıfın öğrencileri veya buna benzer etkenler olabilir. Eğer bir araştırmacının farklı sınıfları bağımsız değişkenin farklı düzeylerine yansız ataması gerekseydi, bütün hâlinde grupları test etmenin neden olduğu bir karıştırıcı etken ortaya çıkabilirdi.

Katılımcılar bütün hâlinde gruplar arasında sistematik şekilde farklılaştığı zaman, denk olmayan grupların neden olduğu bozucu faktör ortaya çıkar. Örneğin, sabah 08.00 dersini seçen öğrenciler, saat 14.00 dersini seçen öğrencilerden farklılaşabilir. Bu bütün hâlinde grupların deneysel koşullara seçkisiz atanması, bu gruplar arasındaki sistematik farklılıkları dengelemeye yetmez. Bu problemin çözümü basittir –bir seçkisiz gruplar deseninde bütün hâlinde gruplar kullanmayın.

Dış Değişkenleri (Extraneous Variables) Dengelemek Araştırma yürütülürken, bir deneydeki çok sayıda faktör, pratik nedenlerle değişir. Örneğin, bir araştırmacı bir deneyi daha çabuk tamamlamak için, farklı araştırmacıların küçük katılımcı gruplarını test etmesine karar verebilir. Bu grupların büyüklükleri ve araştırmacıların kendileri, deney için potansiyel karıştırıcı değişkenler olabilir. Örneğin, deney grubundaki tüm denekleri bir araştırmacı, kontrol grubundakileri ise bir başka araştırmacı test ederse, deneycinin kimliği bir karıştırıcı değişken (confounding variable) hâline gelir. Gruplar arasında gözlenecek farkın bağımsız değişkene mi yoksa deney ve kontrol gruplarındaki katılımcıları farklı araştırmacıların test etmesine mi bağlı olduğunu anlayamayız.

Araştırmacının doğrudan ilgi alanında olmayan, fakat karışıklık kaynağı olabilecek potansiyel değişkenlere *dış değişkenler* denir. Fakat bu terimin sizi yanıltmasına izin vermeyin. Dış değişkenin deney ortamında yarattığı karıştırıcı etki, önemli bir içsel ilginin yarattığı karıştırıcı etkiden daha az değildir. Örneğin, akademik dönem başlangıcında gönüllü olan öğrenciler, daha geç gönüllü olan öğrencilere kıyasla daha akademik yönelimlidir ve iç odak kontrollü olmaları daha olasıdır (yani, davranışlarından dış etkenlerin değil, kendilerinin sorunlu olduğunu düşünmeye eğilimlidirler) (Evans & Donnerstein, 1974). Eğer deneysel koşuldaki tüm katılımcılar dönem başında ve kontrol koşulundaki katılımcılar dönem sonunda test edilseydi, bağımsız değişken katılımcıların özellikleri ile karışacaktı.

Blok seçkisizleştirme dış değişkenleri gruplar arasında dengeleyerek kontrol eder. Gerekli olan tek şey, tüm blokların ikincil değişkenlerin her düzeyinde test edilmesidir. Örneğin, dört farklı deneyci olsaydı, her bir deneyci, blokların tüm koşullarına atanırdı. Çünkü her blok deneyin tüm koşullarını içermektedir, böylece her araştırmacının her koşulu test etmesi garantilenmiş olur. Daima her araştırmacıya aynı sayıda blok tahsis edilir, fakat bu zorunlu değildir. Zorunlu olan şey, tüm blokların ikincil değişkenin her düzeyinde

test edilmesidir ki bu durumda ikincil değişken düzeylerini dört araştırmacı temsil etmektedir. Birçok ikincil değişken var ise, dengeleyici hareket biraz aldatıcı olabilir, fakat dikkatli bir planlama bu tür etkenlerin yaratacağı karıştırıcı etkileri önleyebilir.

Denek Kaybı Seçkisiz gruplar deseninin mantığı grupların deneyde sadece bağımsız değişkenin düzeyleri açısından farklılaşmasını gerektirir. Bir deneyin başlangıcında denk katılımcı gruplar oluşturmak, seçkisiz gruplar deseninin temel özelliğidir. Deney sonunda grupların, bağımsız değişken dışında denk olması da aynı şekilde önemlidir. Denekler deneye başlar, fakat başarılı bir şekilde tamamlayamazlarsa, deneyin içsel geçerliği zayıflar. Öğrencilerin deneyi tamamlamakta başarısız oldukları iki ayrı duruma bakalım: Mekanik denek kaybı ve seçici denek kaybı

Temel Kavram

Mekanik denek kaybı, denekler donanım yetersizliği nedeni ile deneyi tamamlamakta başarısız olduğunda ortaya çıkar (bu durumda araştırmacının donanımın bir parçası olduğu göz önünde tutulmalıdır). Eğer bilgisayar çökerse ya da kendisine yanlış yönerge okunursa veyahut birisi kazara deneysel oturumu keserse, mekanik denek kaybı ortaya çıkar. Mekanik denek kaybı seçici denek kaybından daha az önemli bir problemdir, çünkü bu kayıp deneğin herhangi bir özelliği ile ilişkili değildir. Bu nedenle, mekanik kayıp deneydeki farklı deneysel koşullarda deneyi başarılı bir şekilde tamamlayan deneklerin özellikleri açısından sistemli farklara yol açmamalıdır. Mekanik kayıp aynı zamanda mantıklı olarak gruplar arasında eşit biçimde ortaya çıkması beklenen rastlantısal olayların bir sonucu olarak da anlaşılabilir. Bu yüzden, deneklerin mekanik kayba bağlı olarak deneyden çıkarılması, iç geçerliği zayıflatmaz. Mekanik denek kaybı ortaya çıkarsa, kaydedilmelidir. Deneğin ismi ya da çıkan deneğin denek numarası ve kaybın nedeni kaydedilmelidir. Kaybedilen deneğin yerine, bir sonraki denekle devam edilir.

Temel Kavram

Seçici denek kaybı çok daha ciddi bir durumdur. **Seçici denek kaybının** ortaya çıktığı durumlar şunlardır: (1) Denekler deneysel koşulların her birisinde farklı nedenlerle kaybedildiği zaman; (2) Bu kaybın nedeni, deneğin bazı özellikleri olduğu zaman ve (3) Bu denek özelliği bağımlı değişkenle ilişkili olduğu zaman. Seçici denek kaybı, seçkisiz gruplar deseninin temeli olan grupların denkliğini ortadan kaldırır ve bu nedenle deneyi yorumlanamaz hâle getirir.

Seçici denek kaybı ile ilgili kurmaca fakat gerçekçi problemleri gösterebiliriz. Varsayın ki bir fitnes merkezinin yöneticileri 1 aylık fitnes programının etkinliğini test etmeye karar verdiler. Deneye 80 kişi gönüllü katılıyor ve iki gruptan birine seçkisiz olarak atanıyor. Seçkisiz atama yöntemiyle deneyin başında ağırlık, bedensel görünüş, motivasyon ve benzeri denek özellikleri iki grup arasında dengeleniyor. Kontrol grubu üyelerinden sadece ay sonunda fitnes testini almaları isteniyor. Deneysel gruptaki katılımcılar test öncesinde 1 ay boyunca etkin programa katılıyorlar. Varsayın ki kontrol grubundaki 38 katılımcı ay sonunda fitnes testine geldiler, fakat deneysel koşuldaki katılımcıların sadece 25'i tüm ay boyunca bu etkin programa katıldı. Yine varsayın ki, deneysel gruptaki

İLERİ ALIŞTIRMA II

Bu alıştırmada bir deste oyun kartına ihtiyacınız olacak. Kartları bir tarafa yüzleri açık olarak yerleştirin (Vale, Dam, Papaz) ve birden ona kadar dizin (As, bir değerini alsın). Kâğıtları iyi karıştırın.

Denk gruplar oluşturmak için seçkisiz atamanın nasıl işlediğine dair bir fikir edinmek için, iyice karıştırılmış (seçkisizleştirilmiş) kâğıtları, 20'şerlik iki desteye ayırın. Bu iki destenin birisi deneklerin seçkisiz atandığı deney koşulunu, diğeri ise yine deneklerin seçkisiz atandığı kontrol koşulunu temsil etmektedir. Her bir kartın üzerindeki değer, katılımcının örneğin, bellek yeteneği testinden aldığı 1-10 arasında değişen puanını gösteriyor olsun.

1 Her bir koşuldaki (destedeki) katılımcılar için her karttaki değeri toplayın ve 20'ye bölerek ortalama puanı hesaplayın. İki grup ortalama bellek yetenekleri açısından denk midir?

Seçici denek kaybı ile ilişkili problemleri anlamak için, varsayın ki düşük bellek yeteneği olan katılımcılar (1-2 değeri almış olsun) bir deneysel görevi tamamlayamamış ve deneysel koşuldan çıkarılmış olsunlar. Bu durumu canlandırmak için 1 ve 2 değerli kartları deneysel koşulunuza denk gelen desteden çıkarın.

2. Deneysel koşulu oluşturan deste için yeni bir ortalama puan hesaplayın. Seçici denek kaybından sonra, iki grubun ortalama bellek yeteneği puanları nasıl karşılaştırılabilir? Seçkisiz atamayı kullanarak başlangıçta oluşturulan iki grubun denkliği açısından bu ne anlama gelir?
3. Deneysel gruptan çıkarılan her katılımcı (kart) için, kontrol koşulundan bir kart çıkarın. İki gruptan çıkan kartların değerlerinin aynı olmayacağını, ve "1"in "2" ya da bir başka kartın yerine geçebileceğini unutmayın. Kontrol grubu için de yeni bir ortalama hesaplayın. Bu işlem, iki grubun başlangıçtaki denkliğini tekrar sağlar mı?
4. Kartları tekrar karıştırın ve bu defa 10'ar kartlık dört gruba ayırın. On kartlık bir desteden yeni bir ortalama hesaplayın. Her grupta daha az sayıda "katılımcı" (kart) ile seçkisizleştirme denk gruplar oluşturur mu?

25 kişinin ortalama fitnes puanı, kontrol grubundaki katılımcıların ortalama puanından anlamlı düzeyde yüksek olsun. Fitnes merkezi yöneticileri bu durumda "Bilimsel temelli bir araştırma göstermiştir ki, bizim programımız daha iyi bir forma yol açmaktadır." iddiasında bulunuyorlar.

Fitnes merkezinin bu iddiada haklı olduğunu düşünür müsünüz? Cevap "hayır"dır. Bu hipotetik çalışma seçici denek kaybı için klasik bir örnektir; bu yüzden, çalışmanın sonuçları fitnes merkezi sonuçlarını desteklemek için kullanılamaz. Deney ve kontrol koşullarındaki denek kaybı aynı değildir; kaybın büyük kısmı deneysel gruptandır. Farklı biçimde kayıp olması problemi. grupların farklı büyüklüklere düşmesinden kaynaklanmamaktadır. Gruplardan farklı oranlarda denek kaybı olmasının neden olacağı sorun, grupların farklı sayıda denekle deneyi tamamlamaları değildir. Eğer 25 kişi seçkisiz olarak deneysel koşula ve 38 kişi de kontrol grubuna atanmış olsa ve hepsi deneyi tamamlamış olsa, sonuçlar geçerli şekilde karşılaştırılabilir. Daha ziyade, seçici denek kaybı bir problemdir çünkü deney grubunda programı tamamlayan 25 kişinin, kontrol grubundaki 38 kişiye denk olması mümkün değildir. Deney grubunda programı tamamlayamayan 15 katılımcının, deneyden önce bile, programı tamamlayan 25 kişiden daha az fit oldukları düşünülebilir. Deney grubundaki katılımcıların kaybı, deneyin başında seçkisiz yöntemle oluşturulan grup denkliğini bozmuştur. Aslında, deney grubundaki 25 katılımcının program sonundaki fitnes puanları, deney öncesinde de kontrol grubunun puan ortalamasından yüksek olmuş olabilir, çünkü zaten bu katılımcıların bedenleri programa başlamadan önce de diğerlerine göre daha fitti.

ŞEKİL 6.3 Yorucu bir alıştırma programına başlayan pek çok kişi programı tamamlayamaz. Bir anlamda, eğer farklı fitnes programları karşılaştırılıyor olsaydı, sadece en fit olanlar programı tamamlayabilecekti ki bu durum yorumlama sorunlarına yol açardı.



Şöyle ki, söz konusu kayıp, deneklerin programdan önceki fitnes farklılıklarından kaynaklanıyor olabilir ve bu farklılıklar, araştırmanın sonuçlarıyla bağlantılı olabilir (bk. Şekil 6.3).

Eğer seçici denek kaybı deneyin sonuna kadar devam etmişse, yorumlama sorunu olan bir deney yapmış olmanın kazandırdığı deneyimin dışında, bir katkısının olmadığını söyleyebiliriz. Araştırmacılar, seçici denek kaybının araştırmanın ilerleyen aşamalarında da sorun olabileceğini fark ettikleri zaman önlem almalıdırlar. Bu konuda bir alternatif, bir ön test yapmak ve kaybedilme olasılığı olan denekleri elemektir. Örneğin, alıştırma aşamasında, bir başlangıç fitnes testi verilebilirdi ve böylece sadece belirli bir düzeyin üstünde puan alanlar deneye alınabilirdi. Katılımcıları bu şekilde bir ön elemeden geçirmenin potansiyel bir maliyeti olacaktır. Bu durumda çalışmanın sonuçları muhtemelen, sadece asgari fitnes düzeyinin üzerindeki insanlara genellenebilecekti (dış geçerlik meselesi). Ancak bu maliyete fazlası ile değebilir, çünkü yorumlanabilir bir sınırlı geçerlik, yorumlanma sorunu olan bir araştırmadan üstündür.

Araştırmacılar seçici denek kaybı olasılığı ile karşılaştıkları zaman, bu durumu önlemek için kullanabilecekleri ikinci bir yaklaşım daha vardır. Tüm deneklere, koşullara atanmalarından önce araştırmacılar tarafından bir ön test verilebilir. Bu durumda, eğer deneysel gruptan denek kaybı olursa, benzer ön test puanına sahip bir denek de kontrol grubundan çıkarılabilir. Bu durumda

bu yaklaşım, grupların başlangıçtaki karşılaştırılabilirliğini tekrar yapılandırmaya çalışır. Araştırmacının seçici denek kaybına yol açabilecek olası etkenlerin farkında olması ve uyguladığı ön testin bu etkenleri ölçtüğünden emin olması önemlidir.

Talep Özellikleri ve Deneyci Etkileri Betimleyeceğimiz iç geçerlik açısından bu son zorlayıcı durum hem katılımcı hem de araştırmacının beklentilerinden kaynaklanmaktadır. Talep özellikleri, katılımcıların beklentilerinden kaynaklanan olası yanlılık demektir (Orne, 1962). *Talep özelliği*, bir psikoloji çalışmasında katılımcıların davranışlarını yönlendirmek için kullandıkları ipuçlarına ve diğer bilgilere karşılık gelir (bk. Bölüm 4). Örneğin, bir araştırmaya katılan katılımcılar kendilerine deney sırasında alkol verileceğine dair bilgilendirilmiş ise, gevşeme ve baş dönmesi gibi belli etkileri yaşama beklentisi içerisinde olurlar. Bu durumda, alkolün onlarda tek başına yarattığı etkiden ziyade, kendi beklentileri ile uyumlu davranışlar gösterebilirler.

Temel Kavram

Potansiyel yanlılık aynı zamanda araştırmacıların beklentilerinden de kaynaklanabilir. Bu tür yanlılıkları tanımlamak için kullanılan genel terime **deneyci etkisi** denir (Rosenthal, 1963, 1994a). Eğer araştırmacılar farklı gruplardaki katılımcılara, bağımsız değişkenin gerektirdiğinden farklı bir şekilde davranırlarsa, deneyci etkileri ciddi bir karıştırıcı etki kaynağı olabilir. Örneğin alkolle ilişkili olarak yapılan bir deneyde, alkol verilmemiş olanlara kıyasla, verilmiş olan katılımcılara yönergeler daha yavaş okunursa, bu durumda deneyci etkileri ortaya çıkabilir. Uygulama sırasında deneycinin katılımcılara ilişkin gözlemleri yanlı olduğu zaman da deneyci etkileri ortaya çıkabilir. Örneğin, eğer araştırmacılar "içiciler" arasındaki olağandışı motor hareketleri ya da geveleyerek konuşmaları çok daha fazla fark ederlerse, gözleme dayalı yanlılık alkol çalışmasında da ortaya çıkabilirdi (çünkü "içicilerin" bu şekilde davranmalarını "bekliyor" olurlar). (Bölüm 4'teki beklenti etkilerine bakınız.)

Temel Kavram

Araştırmacılar talep özelliğini ve deneyci etkilerine ilişkin problemleri asla tam olarak ortadan kaldıramazlar, ama bu problemlerin kontrol edilmesini sağlayan spesifik araştırma desenleri vardır. Örneğin, talep özelliğini kontrol etmek için **plasebo kontrol grubu** kullanırlar. Plasebo, ilaç ya da başka aktif bir madde gibi görünen, fakat gerçekte aktif olmayan doğal bir maddedir. Bazı araştırmalar plasebo kullanımının, katılımcılarda, bir "ilaç"ın ortaya çıkarması beklenen etkilere benzer tedavi edici etkiler doğurduğunu göstermiştir (örneğin, Kirsch & Sapirstein, 1998). Araştırmacılar tedavinin etkinliğini plasebo ile karşılaştırarak test ederler. Her iki grup da ilaç alma konusunda aynı "farkındalık" içinde olduğu için, ilacın tedavi edici etkisi konusunda benzer beklentilere sahiptir. Bu, talep özelliğinin gruplar için aynı olduğunu gösterir –her iki gruptaki katılımcılar da ilacın etkilerini hissetmeyi beklemektedir. Deneysel gruplar ve plasebo kontrol grubu arasındaki farklardan birisi, katılımcıların ilacı alma konusundaki beklentisine değil, haklı olarak, alınan ilacın gerçek etkisine bağlanabilir.

Temel Kavram

Plasebo kontrol gruplarının çift kör yöntemiyle birlikte kullanılması, hem talep özelliğini hem de deneyci etkilerini kontrol etmeyi sağlar. **Çift kör işlem**inde, hem katılımcı hem de gözlemci, uygulamanın nasıl yapılacağından habersizdir.

Deney sırasında uygulamanın etkinliğini test ederken, çift kör işlemi için iki araştırmacı gerekecektir. İlk araştırmacı ilaç kapsüllerini hazırlayacak ve onları bir şekilde kodlayacaktır; ikinci de her katılımcıya verilen ilaçları tekrar kodlayarak katılımcılara dağıtacaktır. Bu işlem her katılımcının hangi ilacı aldığına ilişkin kesin kayıt sağlar. Bu durumda ne katılımcı ne de ilacı dağıtan (ve ilacın etkilerini gözleyen) araştırmacı, katılımcının hangi tedaviyi aldığını bilmez. Böylece araştırmacının tedaviye ilişkin beklentileri kontrol edilmiş olur, çünkü gözlemleri yapan araştırmacı kimin geçekten tedavi edici ilacı, kimin plasebo aldığından haberdar değildir. Yine bu şekilde, talep özellikleri de kontrol edilmiş olur, çünkü katılımcılar ilaç mı yoksa plasebo mu aldıklarından habersizdirler.

Plasebo kullanan kontrol grupları, bir taraftan talep özelliklerini kontrol ederken, aynı zamanda tedavi etkinliğini değerlendirmek için değerli bir araştırma aracıdır. Böyle olmakla birlikte, plasebo kontrol gruplarının kullanımı özel etik meseleler ortaya çıkarır. Plasebo kontrol grupları kullanılarak kazanılan bilginin faydaları değerlendirilirken, ilaç aldığını zannedip plasebo alan katılımcılar açısından ortaya çıkabilecek riskler göz önüne alınmalıdır. Genel bir uygulama olarak, uygulamanın etik olması katılımcılardan, deney başlamadan önce bilgilendirilmiş onam alınması ile sağlanabilir. Bu onamda katılımcılara, ilaç ya da plasebo alabilecekleri söylenir. Araştırmaya sadece bunu kabul eden katılımcılar katılabilir. Deneysel olarak ilacın etkinliği kanıtlanırsa, deneyden sonra plasebo koşulundaki katılımcılara ilaçla tedavinin de önerilmesi etik açıdan, gereklidir.

DENEYSEL BULGULARIN ANALİZİ VE YORUMLANMASI

Deneylerde Veri Analizinin Rolü

- Veri analizi ve istatistik, bağımsız değişkenin bağımlı değişken üzerinde etkili olduğu iddiasında önemli bir role sahiptir.
- Bir deneyin bulgularının güvenilir olup olmadığını belirlemenin en iyi yolu, o deneyi tekrar yapmaktır (replication). İyi bir araştırma, daima iyi bir araştırma sorusu ile başlar.

Araştırmacılar daha sonra, araştırmanın sorularına yorumlanabilir kanıtlar toplanmasını mümkün kılacak bir deney tasarlamak ve uygulamak için kontrol teknikleri kullanırlar. Bununla birlikte, sadece iyi bir deney yapmak yeterli değildir. Araştırmacılar aynı zamanda bulgularının, vardıkları sonuçları desteklediğini ikna edici bir şekilde göstermelidir. Veri analizi ve istatistiksel analiz, deneysel kanıtların analizinde ve yorumlanmasında önemli bir rol oynar.

Robert Abelson, *Statistics as Principled Argument* (1995) adlı kitabı'nda, veri analizinin temel amacının, gözlemlerin davranış hakkındaki bir iddiayı desteklemekte olup olmadığının belirlenmesi olduğunu öne sürer. Bu durumda, bir deneyde toplanan kanıta dayalı bir sonuç geliştirebilir miyiz?

Araştırmacıların veri analizini ve istatistiği nasıl kullandıkları Bölüm 11 ve 12'de ayrıntılı şekilde açıklanmıştır. Burada deney sonuçlarını yorumlamak için uygulanan veri analizindeki temel kavramları ele alacağız. Fakat ilk önce, araştırmacıların kendi çalışmalarının sonuçlarını ortaya koyabilmelerini sağlayan önemli bir yöntemden bahsedelim.

Temel Kavram

Bir deneyde elde edilen bulguların güvenilir (tutarlı) olup olmadığını belirlemenin en iyi yolu, deneyi tekrarlamak ve aynı soncun elde edilip edilmediğini görmektir. **Tekrar çalışması** (replication), bir deney ikinci kez yapıldığında, aynı sonuçların elde edilip edilmediğini görmek üzere, o deneyde daha önce uygulanmış olan tüm işlemlerin tekrarlanması demektir. Tasavvur edebileceğiniz gibi, tekrarın tam tamına gerçekleştirilebilmesi imkansızdır. Tekrar çalışması sırasında denekler, orijinal araştırmada test edilen katılımcılardan farklı olacaktır; aynı zamanda test odaları ve araştırmacılar da farklı olabilir. Yine de tekrar çalışması, bir araştırma bulgusunun güvenilir olup olmadığını belirlemenin en iyi yoludur. Tekrarlama, psikolojik bulguların güvenirliğini göstermenin o kadar önemli bir yöntemidir ki *Perspectives on Psychological Science* adlı çok değerli bir eserde (Pashler & Wagenmakers, 2012) bu konu ayrı ve özel bir başlık altında ele alınmıştır.

Ama bilim insanlarının yaptıkları her araştırmanın güvenirliğini göstermeleri için tekrarlama yapmaları gerekse, bilimin ilerleyişi çok yavaşlar. Deneylerde yer alan katılımcıların tekrar bulunması kolay olmayacaktır ve bir araştırmanın tekrar çalışmasını yapmak, bizim yeni ve farklı sorular soracağımız bir araştırmayı yapmaktan vazgeçmemiz anlamına gelir. İşte veri analizi ve istatistik, tek bir deneyin sonuçlarının güvenilir olup olmadığını ve elde edilen sonucun nedeninin gerçekten de bağımsız değişken olup olmadığını belirlemek için, tekrarlama çalışmasının yerine kullanılır.

İstatistik İpucu

Bir deneyin veri analizi üç aşamadan oluşur: (1) Veriyi tanımaya çalışmak, (2) veriyi özetlemek ve (3) verilerin anlamını onaylamak. İlk aşamada veri grubunda,ne olduğunu anlamaya, hataları bulmaya ve verinin ne anlam ifade ettiğini öğrenmeye çalışırız. İkinci aşamada betimleyici (descriptive) istatistik ve bulunanları özetleyen grafikler kullanırız. Üçüncü aşamada verilerin, davranış hakkında bize ne söylediğine dair kanıt ararız. Bu aşamada çeşitli istatistiksel teknikler kullanarak veri hakkında sonuçlar elde ederiz.

Birbirini izleyen bölümlerde veri analizinin bu üç aşamasından özet olarak bahsettik. Veri analizine ilişkin daha geniş bir açıklama Bölüm 11 ve 12'dedir (Özellikle Kutu 11.1'e bakınız). Bilimsel bir dergide basılmış bir psikoloji deneyinin sonuçlarını okumak ve yorumlamak ya da kendi deneyinizi yürütmek isterseniz, aşağıdaki bölümler özellikle önemlidir.

Veri analizi sürecini, katılımcıların şiddet içeren video oyunları oynarken ödül ve cezanın etkilerini araştıran bir deneyin sonuçlarını ele alarak açıklayacağız. Carnagey ve Anderson (2005) çalışmalarında, çok sayıda araştırmadan elde edilen bulguların, şiddet içeren video oyunları oynamanın, saldırgan duygu

durum, biliş ve davranışı arttırdığını gösterdiğini belirtmiştir. Araştırmacılar, şiddet içeren video oyunları sırasında ortaya çıkan davranışların, pek çok video oyununda olduğu gibi *ödüllendirilmesinin* veya *cezalandırılmasının* fark yaratıp yaratmadığını araştırdılar. Carnagey ve Anderson, oyunu oynayanların davranışları cezalandırılırsa, daha az saldırgan olacaklarını iddia etmiştir. Diğer bir hipotez ise, şiddet içeren davranışları cezalandırılan oyuncuların, engellendikleri için, daha fazla saldırgan olacaklarını öne sürmüştür.

Bu çalışmalarda, lisans öğrencileri laboratuvar ortamı koşullarında, birbiri ile yarışan yarış arabalarının olduğu aynı video oyununun ("Carmageddon 2") üç versiyonundan biri ile oynamıştır. Ödül koşulunda, katılımcılar rakipleri ile yarıştıkları ve yayaları öldürdükleri koşulda ödüllendirildiler (puan aldılar). Cezalandırma koşulunda ise katılımcılar rakiplerini öldürdüklerinde ya da onlara çarptıklarında puan kaybettiler. Üçüncü bir koşulda, oyun şiddet içermeyecek şekilde düzenlendi ve katılımcılar pist etrafında yarışırken, kontrol noktalarını geçtikçe puanlar aldılar (tüm yayalar çıkarılmış ve yarıştaki rakipler pasif olacak şekilde programlanmıştır).

Carnagey ve Anderson (2005) katılımcıların bu üç video oyunundan birine seçkisiz olarak atandığı üç deneyin sonuçlarını açıkladı. Temel bağımlı değişkenler, katılımcıların düşmanca duygulanımlarının ölçümü (Deney 1), saldırganca düşünmeleri (Deney 2) ve saldırgan davranışları (Deney 3) idi. Bu üç çalışma sonucuna göre saldırgan davranışları için ödüllendirilen katılımcılar, cezalandırılan ve şiddet içermeyen koşula kıyasla, daha fazla saldırgan duygulanım, düşünce ve davranışlar gösterdiler. Video oyunundaki saldırgan davranışların cezalandırılması, şiddet içermeyen koşuldakine kıyasla, katılımcıların daha fazla düşmanca duygu (ödüllendirme koşuluna benzer şekilde) yaşamalarına yol açtı. Fakat saldırgan düşünce ve davranış yaşamalarında artışa yol açmadı.

Veri analiz sürecini örneklerle açıklamak için, Carnagey ve Anderson'ın saldırgan düşüncelere ilişkin sonuçlarını daha yakından inceleyeceğiz (Deney 2). Video oyununun bu üç versiyonundan birini oynadıktan sonra, katılımcılara kelime parçalarını tamamlama testi verildi ve beş dakika içerisinde 98 kelimeden tamamlayabilecekleri kadar çok sayıda kelimeyi tamamlamaları istendi. Kelime parçalarının yarısı saldırgan olasılıklar içermekteydi. Örneğin, bir kelime parçası "K I ___ ___", "kiss" (öpmek) ya da "kill" (öldürmek) ile tamamlanabilirdi.

Saldırganlığın işevuruk ölçütü, bir katılımcının saldırgan kelimelerle (örneğin, "öldürmek") tamamladığı kelime parçaları sayısıydı. Örneğin, eğer bir katılımcı kelime parçalarının 60'ını beş dakikada tamamladıysa ve bunların 12'si saldırgan içerikliyse, saldırgan koşul puanı .20 (12/60 =0.20) oldu.

Sonuçların Betimlenmesi

- Deneylerin sonuçlarını özetlemek için kullanılan en yaygın iki istatistik, ortalama ve standart sapmadır.
- Etki büyüklüğü ölçümleri, bağımlı ve bağımsız değişken arasındaki ilişkinin gücünü gösterir ve örneklem büyüklüğünden etkilenmezler.

- Etki büyüklüğünün yaygın kullanılan bir ölçümü olan *d*, iki grubun ortalamalarının, deneydeki ortalama değişkenliğe göre farkını inceler.
- Meta-analiz aynı bağımsız ve bağımlı değişkeni araştıran bir çok deneyin sonuçlarını özetlemek için etki büyüklüğü ölçümlerini kullanır.

Veri analizi özellikle, veriler arasında olası hataların ya da anormal puanların olup olmadığını anlamak amacıyla, tüm verilerin dikkatli bir şekilde incelenmesiyle başlamalıdır. Veriyi incelemek için kullanılan teknikler Bölüm 11'de betimlenmiştir. Sonraki aşama, ne bulunduğunu belirlemektir. Bu aşamada araştırmacı "deneyde ne olduğunu" bilmek ister. Bu soruyu cevaplayabilmek için ise *betimleyici istatistiği* kullanır. En yaygın kullanılan iki betimleyici istatistik, ortalama (bir merkezî eğilim ölçümü) ve standart sapmadır (bir değişkenlik ölçümü). Saldırgan biliş ile ilişkili video oyunu deneyinde ortalama ve standart sapmalar Tablo 6.1'de sunulmuştur. Ortalamalar saldırgan bilişin ödül koşulunda en yüksek (.210) ve şiddet olmayan koşulda en düşük (.157) olduğunu göstermektedir. Ceza koşulundaki saldırgan biliş (.175), şiddet uygulanmayan ve ödül koşulları arasındadır. Ödül koşulundaki katılımcıların tamamladıkları kelime parçalarının yaklaşık 5'te 1'inin saldırgan içerikli kelimelerle tamamlandığını söyleyebiliriz (zaten kelime parçalarının ancak yarısı saldırgan içerikli olacak şekilde tamamlanabiliyordu).

Uygun şekilde yürütülen bir deneyde, her gruptaki standart sapma, sadece o gruba seçkisiz atanan denekler arasındaki bireysel farklılıkları ifade etmelidir. Her gruptaki denekler aynı şekilde işlem görmeli ve atandıkları bağımsız değişken düzeyi, gruptaki her denek için aynı şekilde gerçekleştirilmelidir. Tablo 6.1'deki standart sapmalar, değişkenliğin her grupta ortalamalar etrafında olduğunu ve üç grubun tümünde de aynı olduğunu göstermektedir.

Temel Kavram

Bir deneyin sonuçlarını anlatırken araştırmacıların sorduğu önemli bir soru, bir bağımsız değişkenin bağımlı değişken üzerindeki etkisinin ne kadar büyük olacağıdır. **Etki büyüklüğü** ölçümleri, bu soruyu cevaplamada kullanılabilir, çünkü bağımsız ve bağımlı değişkenler arasındaki ilişkinin gücünü gösterir. Etki büyüklüğü ölçümlerinin bir avantajı da deneyde test edilen örneklemlerin büyüklüğünden etkilenmemeleridir. Etki büyüklüğü ölçümleri, bir deneydeki iki koşulun ortalamaları arasındaki farktan daha fazla dikkate alınır.

TABLO 6.1 VİDEO OYUNU DENEYİNİN ÜÇ KOŞULUNDA ORTALAMA SALDIRGAN BİLİŞ, STANDART SAPMALARI VE GÜVEN ARALIKLARI

Video Oyun Versiyonu	Ortalama	*S*	.95 Güven Aralığı*
Ödül	.210	.066	.186–.234
Ceza	.175	.046	.151–.199
Şiddet - yok	.157	.050	.133–.181

*Güven aralıkları Carnagey ve Anderson'ın (2005) bildirimleri temel alınarak hesaplanmıştır.

Temel Kavram

Gruplar arasındaki ortalama fark, daima katılımcıların puanlarındaki ortalama değişkenliğe göredir. Etki büyüklüğünün sıklıkla kullanılan bir ölçümü **Cohen's *d***'dir. Cohen (1992) günümüzde yaygın şekilde kabul gören işlemler geliştirdi. Ona göre, .20, .50, ve .80 d değerleri sırasıyla, bağımsız değişkenin bağımlı değişken üzerindeki etkisinin küçük, orta ve büyük olduğu anlamına gelir.

Video oyunu deneyindeki iki koşulu, ödül ve şiddet içermeyen koşulları, kıyaslayarak, Cohen *d*' nin etki ölçümü olarak nasıl kullanıldığına örnek verebiliriz. Ödül koşulundaki (.210) ve şiddet içermeyen koşuldaki (.157) saldırgan biliş arasındaki farka ait *d* değeri .83'tür. Bu *d* değeri, bize şunu söyleme olanağı verir: Ödül ve şiddet-yok koşullarını içeren bağımsız değişken, bu iki koşuldaki saldırgan bilişi (bağımlı değişkeni) yüksek düzeyde etkilemektedir. Bu etki büyüklüğü ölçümleri, deneyin bulgularını açıklamada araştırmacılara değerli bilgiler sağlar.

İstatistik İpucu

Bölüm 11 ve 12'de, etki büyüklüğünün yanı sıra, merkezî eğilim ve değişkenlik ölçümleri de açıklanmaktadır. Bu bölümlerde bu ölçümler için ana hatlarıyla hesaplama aşamalarını açıklıyor ve nasıl yorumlanacaklarını tartışıyoruz. Psikoloji literatüründe, bir çok farklı etki büyüklüğü ölçümü vardır. Cohen's d dışında bir başka popüler etki büyüklüğü ölçümü, bağımsız ve bağımlı değişken arasındaki ilişkinin gücünü ölçen, eta karedir (Bölüm 12'ye bakınız). Eta kare bağımsız değişkenin, bağımlı değişken üzerindeki değişkenliğin ne kadarını açıklayabildiğini söyler. Etki büyüklüğü ölçümlerinin kullanımı iki, üç ya da daha çok çalışmadan elde edilen sayısal değerleri karşılaştırırken veya meta-analizde olduğu gibi (aşağıya bakınız), araştırmalar arasında ölçümlerin ortalamasını alırken çok yararlıdır.

Temel Kavram

Araştırmacılar etki büyüklüğü ölçümlerini meta-analiz adı verilen bir işlemde de kullanırlar. **Meta-analiz,** aynı bağımsız ya da bağımlı değişkeni araştıran çok sayıda ayrı deneydeki etki büyüklüğünü özetlemek için kullanılır. Genel olarak diyebiliriz ki bir meta-analizin değerini, analiz ettiği deneylerin yöntemsel değeri belirler (bk. Judd, Smith & Kidder, 1991). Meta-analizler, aynı probleme yönelik soruları cevaplamak için kullanılır: Uyumluluk (conformity) cinsiyete göre değişir mi? Sınıfın büyüklüğü, akademik başarıyı etkiler mi? Bilişsel tedavi yöntemi, depresyonun tedavisinde etkili midir? Kutu 6.2'de psikolojik sorunları olan gençlere yönelik psikoterapi ile ilişkili meta-analiz çalışmalarını göreceksiniz. Tek tek deney sonuçları, bu deneyler ne kadar iyi yapılmış olursa olsun, çoğunlukla bu gibi önemli genel meseleleri ele alan soruları cevaplamakta yetersiz kalır. Her mesele ile ilişkili olarak birçok araştırmayı (deneyi) ele almamız gerekir (meta-analize iyi ve anlaşılır bir giriş yapmak için Hunt, 1997'e bakınız). Meta-analiz sonuçları, ayrı ayrı yapılmış çok sayıda deneyin sonuçlarının birlikte değerlendirilmesi ile ortaya çıkar. Bu nedenle meta-analiz, psikolojinin prensipleri hakkında daha güçlü sonuçlar çıkarmamıza olanak sağlar. Bu analizler etki büyüklüğü ölçümlerini kullanarak çok sayıda çalışmanın sonuçlarını özetlemek için verimli ve etkili bir yoldur.

KUTU 6.2

BİR META-ANALİZ ÖRNEĞİ "OLAĞAN KLİNİK TEDAVİYE KARŞI, KANITA DAYALI GENÇLİK PSİKOTERAPİLERİ"

Weisz, Jensen-Doss ve Hawley (2006) gençlerde, kanıta dayalı tedavi (evidence-based treatments) ile klasik tedavinin (usual care) etkilerini karşılaştırmak için, 32 psikoterapi araştırmasını içeren bir meta-analiz yaptılar. Kanıta dayalı tedavi (KDT), bireylere yardım etmek için kendilerine görgül (ampirik) destek verilen bir uygulamadır. Bu gözleme dayalı destek nedeniyle, KDT'nin klinik uygulamada yaygın kullanılmasının gerekliliği açık olsa da, bir çok terapist bu uygulamaların, alışılagelmiş klinik durumlarda etkin olamayacağını öne sürmektedir. Bazı psikologlar KDT'nin, danışanların kendilerine özgü ihtiyaçlarına göre şekillendirilemeyen, sert ve esnek olmayan bir yöntem olduğunu iddia etmişlerdir. Dahası, KDT'ye karşı çıkanlar, bu uygulamanın etkinliğini gösteren araştırmaların, klasik tedaviye göre çok daha basit vakaları ele alan görgül çalışmalara dayandıklarını öne sürmüşlerdir. Bu tartışmalar şunu düşündürmektedir: Psikoterapi, danışmanlık ya da zihinsel sağlık hizmeti verenlerin düzenli şekilde yaptıkları vaka yönetimi gibi klasik tedaviler (KT), danışanların ihtiyaçlarını daha iyi karşılamaktadır.

Weisz ve arkadaşları, KDT ile KT yöntemlerine ait sonuçları doğrudan karşılaştırmak için meta-analiz yöntemini kullandılar. KDT ve KT'yi karşılaştıran 32 çalışma arasında, ortalama etki büyüklüğü 0.30'du. Böylece, KDT ile tedavi edilen gençlik, KT uygulanan gruba kıyasla ortalama açısından daha iyi durumdaydı. Cohen's d (1988) kriterlerine göre .30 değeri, küçük ve orta düzey etkilerin arasına düşer. Bu etki büyüklüğü iki tip tedavi arasındaki farkı temsil eder; psikoterapinin kendisinin etkisini değil. Weisz ve arkadaşları, KDT uygulananların, hiç bir uygulama yapılmayan kontrol grupları ile karşılaştırıldıklarında, KDT etki büyüklüğünün tipik olarak 0.50 ile 0.80 aralığında yer aldığını bildirmişlerdir (orta ve büyük etki düzeyleri). Ek analizlerde araştırmacılar, bu 32 araştırmayı tedavi edilen problemlerin şiddet ve karmaşıklığı, tedavi ortamı ve terapistlerin özellikleri gibi özelliklerine göre grupladılar. Bu analizler, KDT'ye yöneltilen eleştirilerin, KT'nin uygulanmasına devam edilmesini haklı çıkaracak nitelikte olup olmadığını belirlemek için yapıldı. Weisz ve arkadaşları, çeşitli faktörlere göre çalışmaları bu şekilde gruplamanın, KDT'nin KT'den daha iyi bir uygulama olduğuna ilişkin sonuçları etkilemediğini buldu.

Bu meta-analiz psikologların, psikoterapiye ilişkin genel bir psikolojik ilkeyi daha büyük bir güvenle öne sürmelerine izin verir: Kanıta dayalı tedaviler, klasik tedaviye kıyasla, gençlik için daha iyi sonuçlar veriyor.

Verilerin Anlamının Onaylanması

- Araştırmacılar bir bağımsız değişkenin bağımlı değişken üzerinde güvenilir bir etkisi olup olmadığını belirlemek için çıkarımsal istatistik kullanırlar.
- Örneklemden elde edilen verilere dayanan çıkarımlar yapmak için kullanılan iki yöntem, sıfır hipotezinin test edilmesi ve güven aralığıdır.
- Araştırmacılar, bir deneyde gruplar arasındaki ortalama farkların, sadece hata değişkenliği nedeniyle beklenen farklardan daha fazla olup olmadığını belirlemek için sıfır hipotezini test ederler.
- Eğer sıfır hipotezi doğru ise, istatistiksel olarak anlamlı bir sonuç ortaya çıkma olasılığı düşüktür.
- Araştırmacılar bir deneydeki farklı örneklemlere ilişkin güven aralıklarının örtüşüp örtüşmediğini inceleyerek, bir bağımsız değişkenin davranış

üzerinde etkisi olup olmadığını belirlerler. Örtüşme derecesi, örneklem ortalamalarının aynı evrenin ortalamasını mı yoksa farklı evrenlerin ortalamalarını mı tahmin ettiğini anlamamızı sağlar.

Belki de araştırmacıların deney yaparken öne sürmek istedikleri en temel iddia, bağımsız değişkenin bağımlı değişken üzerinde etkisinin olduğu iddiasıdır. Bu iddiayı ifade etmenin bir diğer yolu, araştırmacıların bağımsız değişkenin *davranış üzerinde bir farklılık yarattığını* doğrulama istekleridir. Betimleyici istatistik, bu temel iddiayı doğrulamak için yeterli değildir.

Bağımsız değişkenin deneyde bir etkiye yol açtığını doğrulamak için, araştırmacılar *çıkarımsal istatistik* kullanırlar. Seçkisiz atama yolu ile sağlanan kontrolün doğası gereği, çıkarımsal istatistiğe ihtiyaç duyarlar. Daha önce belirttiğimiz gibi, seçkisiz atama denekler arasındaki bireysel farkları *ortadan kaldırmaz*. Seçkisiz atama, deneydeki gruplarda yer alan denekler arasındaki bireysel farkları basitçe *dengeler* ya da ortalamasını bulur. Her gruptaki denekler arasında mevcut farklara bağlı sistematik olmayan (yani, seçkisiz) değişkenliğe *hata varyasyonu* denir. Hata değişkenliğinin ortaya çıkışı potansiyel bir problem içerir, çünkü deneyde bulunan grup ortalamaları arasındaki farkın nedeninin bağımsız değişken değil, basitçe hata değişkenliği olduğu anlamına gelir. Böylece, çok iyi kontrol edilmiş bir deneyde elde edilen ortalamalar bile, kendi başına bağımsız değişkenin davranış üzerinde bir farklılık üretip üretmediği hakkında belirli bir sonuca varılmasını engeller. İstatistiksel çıkarım araştırmacıların, grup ortalamaları arasında gözlenen farkların sadece şansa (hata değişkenliğine) mı yoksa bağımsız değişkenin etkisine mi bağlı olduğunu test olanağı verir. Araştırmacılar bir bağımsız değişkenin etkisinin olup olmadığı anlamak için iki tür çıkarımsal istatistik kullanırlar: Sıfır hipotezini test etme ve güven aralıkları.



En iyi kontrol edilmiş deneyin sonuçları bile, bir bağımsız değişkenin davranış üzerinde bir değişikliğe yol açıp açmadığı konusunda belirli bir sonuca varmaya izin vermeyebilir; bunu öğrenmenin hayal kırıklığı yaratacağının da farkındayız. Bir başka deyişle, araştırma yöntemleri hakkında bu kadar çok şey öğrenmiş olsanız da yeterli olmadığını görüyorsunuz. Ne yazık ki, veri analizi araçları ile davranıştaki değişikliğin nedeni hakkında *belirli* sonuçlara varamayız. Fakat davranışta ortaya çıkan değişikliğin nedeni hakkında en olası tahmini yapmanız için bir yol (aslında, bir çok yol) önerebiliriz. Sonuç, *olasılığa* –yani gözlenen etkinin şansa (çoğunlukla hata değişkenliğini ifade eder) bağlı olup olmadığına karar vermenize yardım edecek bir olasılığa– dayanır. Sıfır hipotezini test etmenin karmaşasında ve güven aralıklarında kaybolmak kolaydır, fakat şu iki kritik noktayı aklınızda tutun:

İlk ve en önemli nokta, davranıştaki farklar şansa (sıklıkla *hata varyasyonuna*) bağlı olarak ortaya çıkabilir. Bilmek istediğiniz şey, bulduğunuz farkın sadece şansa bağlı olmasının, nasıl mümkün olduğu mudur (bağımsız değişkenin etkisi olmaksızın)? Yoksa gerçekte bilmek istediğiniz şey, bağımsız değişkeninizin bir etkisi olmasının nasıl mümkün olduğu mudur? Yine de,

biz bu soruları çıkarımsal istatistik kullanarak cevaplayamayız. İstatistiksel çıkarımın dolaylı olduğunu göreceksiniz (bk. örneğin, Kutu 12.1, Bölüm12).

İkincisi, topladığınız veri evrenden alınmış *örneklemleri* temsil eder; fakat bir anlamda, gerçekte önemli olan örneklemler değil, evrenlerdir (Eğer sadece örneklem ortalamaları önemli olsaydı, o zaman örneklem ortalamalarının farklı olup olmadığına bakabilirdiniz). Deneyinizin farklı koşullarındaki örneklere ilişkin ortalama performans, evrenin ortalamasını *çıkarsamak* için tahminde bulunmanızı sağlar. İstatistiksel çıkarıma dayanan açıklamalar yaptığınızda, evren ortalamaları arasındaki farklar hakkında karar vermek (çıkarım yapmak) için örneklemlerin ortalamalarını kullanırsınız. Bu konulara ilişkin daha ayrıntılı tartışmalar için biz tekrar Bölüm 12'ye bakmanızı öneriyoruz.

Temel Kavram

Sıfır Hipotezi Anlamlılık Testi (SHAT) Araştırmacılar bir bağımsız değişkenin bir deneyde etkiye yol açıp açmadığına karar vermek için çoğunlukla **sıfır hipotezi anlamlılık testi (SHAT)** kullanırlar. Sıfır hipotezinin anlamlılığını test etme, bağımsız değişkenin hiç etkisi olmadığı varsayımı ile başlar. Sıfır hipotezinin doğru olduğunu varsayarsak, deneyimizde gözlediğimiz farkın "şans eseri" ortaya çıkmış olduğuna ilişkin olasılığı belirlemek için, olasılık teorisini kullanabiliriz.

Temel Kavram

***İstatistiksel olarak anlamlı** bir sonuç, eğer sıfır hipotezi doğru olmuş olsaydı, sadece küçük bir ortaya çıkma olasılığı olan sonuçtur.* İstatistiksel olarak anlamlı bir sonuç şöyle de ifade edilebilir: Deneyimizde elde edilen fark, sonucun sadece hata değişkenliğinden (yani, şanstan) kaynaklanması durumunda beklenen farktan daha büyük olması demektir.

Bir deneyin sonucu çoğunlukla, deneydeki koşullara ait ortalamalar arasındaki farklar cinsinden ifade edilir. Bir deneyde elde edilen sonucun ortaya çıkma olasılığını nasıl bilebiliriz? Büyük çoğunlukla araştırmacılar, *t*-testi ya da *F* testi gibi çıkarımsal istatistik testlerini kullanırlar. Bağımsız değişkenin iki düzeyi varsa *t* testi, üç ya da daha fazla düzeyi varsa *F* testini kullanırlar. *t* veya *F* testinin her bir değeri, bu değere karşılık gelen bir olasılığı temsil eder. Araştırmacı bir defa test değerini hesapladıktan sonra, bu olasılık değeri belirlenebilir.

Sonucumuzun olasılık değerinin, istatistiksel olarak anlamlı olması için, ne kadar küçük olması gerekir? Bilim insanları yüzde 5'ten az olasılık (p) değeri olan (ya da $p>.05$) sonuçları istatistiksel olarak anlamlı kabul etme eğilimindedirler. Araştırmacıların bir sonucun istatistiksel olarak anlamlı olduğuna karar vermek için kullandıkları olasılık değerine *anlamlılık düzeyi* denilir. Anlamlılık düzeyi Yunan alfabesindeki harflerden (α) ile gösterilir.

Şimdi daha önce (bk. Tablo 6.1) bahsettiğimiz video oyunu deneyini analiz etmek için kullandığımız sıfır hipotezi işlemlerini örneklerle açıklayabiliriz. Soracağımız ilk araştırma sorusu, video oyunundaki bağımsız değişkenin genel olarak her hangi bir etkisi olup olmadığıdır. Bu durumda saldırgan biliş, bu üç ayrı video oyunu versiyonunda farklılaşmakta mıdır? Testin bütününe ilişkin olarak sıfır hipotezi, deneysel koşul ortalamaları ile gösterilmekte olan (sıfır hipotezinin bağımsız değişkenin hiç bir etkisi olmadığını varsaydığını hatırlayın)

popülasyon ortalamaları arasında hiç bir fark olmadığı anlamına gelir. Video oyunu etkisine ilişkin olarak hesaplanan *F* testi için *p* değeri .05 anlamlılık düzeyinden daha azdı; böylece video oyunu değişkeninin genel etkisi, istatistiksel olarak anlamlıydı. Bu sonucu yorumlamak için,Tablo 6.1'deki bu deney ile ilişkili betimleyici istatistiklere başvurmamız gerekir. Bu durumda üç video oyunu koşulu için saldırgan biliş ortalamasının farklı olduğunu görürüz. Örneğin, saldırgan biliş, ödüllü video oyunu koşulunda en yüksektir (.210) ve şiddet içermeyen koşulda en düşüktür (.157). *F* testinin istatistiksel olarak anlamlı çıkması, bizim video oyunu versiyonunun saldırgan bilişte bir farka yol açtığını öne sürmemize olanak sağlamaktadır.

Araştırmacılar bağımsız değişkenin bir etkisi olduğundan bahsetmek yerine, bağımsız değişkenin davranış üzerindeki etkileri hakkında daha spesifik iddialarda bulunmak isterler. Ortalamalar arasındaki tüm farklara ilişkin *F*-testler, deneyde bir şey olduğunu gösterir, fakat ne olduğu hakkında çok fazla bilgi vermezler. Bağımsız değişkenin etkileri hakkında daha spesifik bilgi kazanmanın bir yolu güven aralıklarını kullanmaktır.

Ortalama Farklarını İncelemek İçin Güven Aralıklarının Kullanımı Video oyunu deneyindeki her üç grup için güven aralıkları Tablo 6.1'de görülmektedir. Bir güven aralığı, evrenin gerçek ortalamasını içerme olasılığının (çoğunlukla .95) ifade eder. Aralık genişliği bizim tahminimizin ne kadar duyarlı olduğunu gösterir (daha dar aralık, daha iyi bir tahmin demektir). **Güven aralıkları** aynı zamanda iki evrenin ortalamaları arasındaki farkı karşılaştırmak için de kullanılır. Biz Tablo 6.1'deki .95 güven aralıklarını, video oyunu versiyonunun saldırgan biliş üzerindeki etkileri hakkında spesifik sorular sormak için kullanabiliriz. Farklı video oyunu gruplarına ilişkin güven aralıklarının örtüşüp örtüşmediğini inceleyerek bunu çözümleriz. *Güven aralıkları örtüşmezse, iki gruba ilişkin evren ortalamalarının farklılaştığından emin oluruz.* Örneğin, ödül koşulunda güven aralığının .186 ile.234 olduğunu gözden kaçırmayın. Bu, .186 ile .234 aralığının ödül koşulundaki saldırgan biliş evren ortalamasını içerme olasılığını .95 olduğunu gösterir (.210 örneklem ortalamasının evren ortalamasının sadece bir *tahmini* olduğunu hatırlayın). Şiddet içermeyen koşulda güven aralığı .133 ile .181'dir. Bu güven aralığı ödül grubuna ilişkin güven aralığı ile hiç örtüşmez (örneğin, şiddet içermeyen koşulun .181 olan üst sınırı, ödül grubunun .186 olan alt sınırından daha düşüktür). Bu kanıtla, ödül koşulundaki saldırgan bilişin, şiddet içermeyen video oyunu koşulundaki saldırgan bilişten daha fazla olduğunu iddia edebiliriz.

Temel Kavram

Ödül grubu (.186–.234) ile ceza grubunun (.151–.199) güven aralıklarını karşılaştırdığımızda, yine de farklı bir sonuca varırız. Bu gruplar için güven aralıkları örtüşür. .210 ve .175 örneklem ortalamaları farklılaşsa da, güven aralıkları örtüştüğü için evren ortalamalarının farklılaştığı sonucuna varamayız. Bu sonucu yorumlamak için şu baş parmak kuralını öneriyoruz: *Eğer aralıklar hafifçe örtüşürse, o zaman doğru ortalama fark konusundaki belirsizliğimizi kabul etmeli ve kararımızı ertelemeliyiz; ama eğer aralıklar bir grup ortalaması bir diğer grubun aralığında yer alacak şekilde örtüşürse, evren ortalamalarının farklılaşmadığı sonucuna varabiliriz.* Video oyunu deneyinde, örtüşme küçüktür ve koşuldaki örneklem

ortalamaları diğer gruba ilişkin aralıklara düşmemektedir. Evrenlerin farklılaşıp farklılaşmadığına karar vermek isteriz, söyleyebileceğimiz tek şey, hangi yolu seçeceğimize dair yeterli kanıtımızın olmadığıdır. Bu durumda kararımızı bir sonraki deney yapılana dek ertelemeliyiz.



Güven aralıkları ile t-testinin mantığı ve hesaplama işlemleri, Bölüm 11'dedir. F-testi ise (farklı şekilleriyle) Bölüm 12'de tartışılmaktadır.

Veri Analizi Bize Ne Söyleyemez

Biz veri analizinin bize söyleyemeyeceği bir şeyden bahsettik. Deneyimizin iç geçerliği olsa ve sonuçlar istatistiksel olarak anlamlı olsa bile, bağımsız değişkenimizin bir etkisi olduğunu (ya da olmadığını) *kesin olarak* söyleyemeyiz. Olasılık ifadeleri ile yaşamayı öğrenmeliyiz. Veri analizimizin sonuçları aynı zamanda çalışmamızın sonuçlarının pratik değeri olup olmadığını hatta anlamlı olup olmadığını da söyleyemez. Önemsiz araştırma soruları soran deneyler yapmak kolaydır (Sternberg 1977, Bölüm 1'e bakınız). Aynı zamanda kötü bir deney yapmak da kolaydır (belki de çok kolay). İç geçerliği olmayan bu kötü deneyler, kolayca istatistiksel olarak anlamlı sonuçlar ve örtüşmeyen güven aralıkları ortaya çıkarırlar; ancak yine de sonuç yorumlanamayacaktır.

Sonuç istatistiksel olarak anlamlı ise, bağımsız değişkenin davranış üzerinde bir etki yarattığı sonucuna varırız. Zaten görüldüğü gibi, "makul bir şüphenin ötesinde" bir sonuca ulaşsak da, analizimiz bize sonucumuz hakkında kesinlik sağlamaz. Aynı zamanda, bir sonuç istatistiksel olarak anlamlı *değilse*, bağımsız değişkenin bir etkisi *olmadığı* kesin sonucuna da varamayız. Tüm varabileceğimiz sonuç, bağımsız değişkenin deneyde bir etki yarattığını öne sürebilmemiz için yeterli kanıt olmadığıdır. Bir bağımsız değişkenin etkisinin olmamış olduğunun belirlenmesi, uygulamalı araştırmalar açısından daha da kritik olabilir. Örneğin, tanınmamış bir ilâç, eşdeğeri olduğu ama iyi bilinen bir ilaç kadar etkili midir?

Bu araştırma sorusuna cevap verebilmek için, araştırmacılar çoğunlukla iki ilaç arasında fark bulmamak için uğraşırlar. Koşullar arasında fark olmadığını göstermeye çalışan deneylerin standartları, bir bağımsız değişkenin etkisinin olduğunu göstermeye çalışan deneylere kıyasla daha yüksektir. Bölüm 12'de bu standartları açıklayacağız.

Araştırmacılar bağımsız değişken etkileri hakkında karar verirken olasılıklara dayandıkları için, her zaman bir hata yapabilirler. Çıkarımsal istatistik kullandıklarında ortaya çıkabilen iki tip hata vardır. Gerçekte sıfır hipotezi (fark yok) doğru olduğu hâlde, bir sonucun istatistiksel olarak anlamlı olduğunu iddia edersek, bu durumda Tip 1 hata yapmış oluruz. *Tip 1 hata*, yangın yokken yangın olduğunu söyleyen yanlış bir alarm gibidir. Gerçekte sıfır hipotezini reddetmek için yeterli kanıtımız olduğu hâlde, olmadığı sonucuna vardığımız durumda *Tip II hata* yapmış oluruz (Tip I ve Tip II hatalar Bölüm

12'de daha ayrıntılı anlatılmıştır.) Sıfır hipotezinin doğru ya da yanlış olduğunu kesin olarak bilebilseydik bu hatalardan her hangi birini yapmayacaktık. Unutmamalıyız ki veri analizi doğru karar vermemizi sağlayabileceği gibi, yanlış karar vermemize de neden olabilir. Araştırmacıların unutmamaları gereken en önemli husus şudur: Çıkarımsal istatistik, bir deneysel bulgunun en son güvenirlik testi gibi, tekrar çalışmasının (replikasyonun) yerine geçemez.

DENEYSEL BULGULARIN DIŞ GEÇERLİĞİNİ SAĞLAMA

- Deney bulguları, spesifik deney alanının dışındaki ortamlara ve koşullara uygulanabildiği zaman, dış geçerlilikleri sağlanır.
- Bazı araştırmalarda (örneğin, teori test etme), araştırmacılar iç geçerliği dış geçerlikten daha önemli görebilirler; başkaları ise örneklemi ve tekrarı (replikasyonu) kullanarak dış geçerliği yükseltmeyi seçebilirler.
- Araştırmacılar için gerçek dünya koşullarında alan deneyleri yapmak, araştırmalarının dış geçerliğini arttırmanın bir yoludur.
- Araştırmacılar çoğunlukla, spesifik koşullar, değişimlemeler, ortam ve örneklemlerden çok, değişkenler arasındaki kavramsal ilişkilere ait sonuçları genellemek isterler.

Bölüm 4'de öğrenmiş olduğunuz gibi *dış geçerlik*, bir araştırmanın bulgularının spesifik bir araştırmanın dışındaki bireylere, ortamlara ve koşullara ne oranda genellenebilir olduğunu ifade eder. Çok iyi kontrol edilmiş olan laboratuvar deneylerine sıkça yöneltilen bir eleştiri, dış geçerlikten yoksun olmasıdır. Yani, kontrollü bir deney ortamında gözlenen bulgular, sadece o özel ortamda olan şeyler (sadece o deneye özgü koşullar ve o deneye katılan katılımcılar) için geçerlidir. Şimdi tekrar, üniversite öğrencilerinin bir laboratuvar ortamında, bir yarış arabası video oyunu oynadıkları deneyi düşünün. Laboratuvar ortamı, bir deneyin kontrol işlemlerinin ideal şekilde hazırlanmasına uygundur ve iç geçerliği en üst düzeye çıkarır. Fakat böyle bir ortamda elde ettiğimiz bulgular, bizim doğal ortamdaki şiddet ve saldırganlığı anlamamıza yardımcı olur mu? Şiddetin farklı şekillerine ne zaman maruz kalınır? Yaşlılar ne zaman şiddete maruz kalır? Bunlar dış geçerlik sorularıdır ve daha genel bir soruyu karşımıza çıkarırlar: Eğer laboratuvarda yapılan deney bulguları sadece o laboratuvar koşullarında geçerli ise, bu bulguların toplum için ne faydası olabilir?

Bu soruya verilecek bir cevap, en azından başlangıçta biraz rahatsız edicidir. Mook'a (1983) göre, bir deneyin amacı psikolojik bir teoriden elde edilmiş spesifik bir hipotezi test etmek olduğu zaman, bulguların dış geçerliğini sorgulamanın önemi kalmaz. Bir deney sıklıkla deneklerin belli bir şekilde davranmak üzere yönlendirilip yönlendirile*meyeceğini* belirlemek için yapılır. Deneklerin doğal çevrelerinde de o şekilde davranıp davranamayacakları meselesi, ikinci derecede önem taşır. Deneylerin dış geçerliliği konusu Riley'nin (1962) aşağıda

ifade ettiği gibi, yeni değildir. "Genel olarak laboratuvar deneyleri doğada var olan en tipik vakayı taklit etmek için düzenlenmezler; araştırmacının ilgi duyduğu bazı soruları cevaplamayı amaçlarlar." (s. 413).

Elbette araştırmacılar, laboratuvar koşullarının dışına genelleyebilecekleri bulgular elde etmek isterler. Bu amaçla, bulgularını genellemek istedikleri koşulların özelliklerini de deney ortamına dahil ederler. Örneğin, Ceci ve arkadaşları (1993), çocuklarda görgü tanıklığı üzerine bir araştırma yaptı. Bu araştırmayı yapmalarının bir nedeninin de önceki çalışmaların *gerçek* bir görgü tanıklığını tüm boyutlarıyla kapsamaması olduğunu bildirdi. Ve yine kendi araştırmasında telkin edici görüşmeler, uzun aralardan sonra hatırlama yaptırma ve rahatsız edici deneyimleri hatırlatma gibi faktörlerin de olduğunu bildirdi. O'na göre, bu faktörlerin araştırmada yer alması, çocuklar tanıklık yaptığında gerçekter ortaya çıkan durumların daha iyi temsil edilmesini sağladı (bk. Şekil 6.4).

Ceci (1993) buna rağmen, yine de deneyler ile gerçek yaşam durumları arasında önemli farkların kaldığına işaret etti:

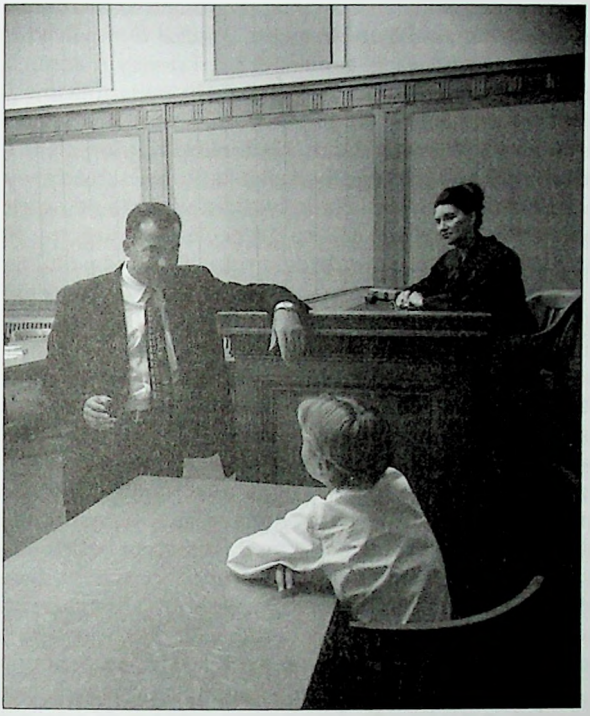
> "Adlî araştırmaları teşvik eden olayların özellikleri, yüksek düzeyde stres, mağdurun bedenine saldırı, ve kontrol kaybıdır. Bu faktörler bizim başka bazı çalışmalarımızda da rol oynamakla birlikte, bir çocuk mağdura yöneltilen saldırgan davranışın saldırgan niteliğini deneysel olarak aynen laboratuvarda yaratamayız, çünkü bu duruma en çok yaklaşan araştırmalar (örneğin, tıbbî araştırmalar) bile, çocuklara yöneltilen cinsel saldırılardan farklı olarak, toplum ve ebeveynler tarafından olumlu karşılanmaz (ss. 41–42)."

Ceci'nin yorumlarının da gösterdiği gibi, bazı durumlarda (örneğin, çirkin olaylar hakkında görgü tanıklığı yapılan durumlarda), deneyin dış geçerliliğini sağlamayı engelleyen önemli etik engeller olabilir.

Araştırma bulgularının dış geçerliği sıklıkla "denekler"in niteliği yüzünden sorgulanır. Farkındaysanız, birçok psikoloji araştırmasında, Psikolojiye Giriş derslerinin bir parçası olarak, üniversite öğrencileri katılır. Dawes (1991), seçili bir grup olması nedeniyle, diğerlerine göre üniversite öğrencilerinin, insan davranışı ve zihinsel süreçleri konusunda genel sonuçlara varmak için her zaman iyi bir örneklem olmayabileceğini belirtmiştir. Benzer şekilde Sue'ye (1999) göre, araştırmacılar dış geçerlikten çok iç geçerliğe önem verdikleri için, katılımcıların evreni temsil etme özelliğine yeterince dikkat edilmemektedir. Araştırılan evrene ilişkin araştırma bulgularının dış geçerliği meselesi, özellikle uygulamalı araştırmalarda önemlidir. Örneğin, tıbbî araştırmalarda, erkekler için etkili bir tedavi kadınlar için etkili olmayabilir ve yetişkinler için etkili olan ise, çocuklar için etkili olmayabilir.

Dördüncü Bölüm'de kısaca belirttiğimiz *alan çalışmaları*, bir araştırmada dış geçerliği arttırmanın bir yoludur. Aynı zamanda pratik bilgi elde edilir. Örneğin, insanların risk algılarını incelemek için, iki alan araştırmasındaki katılımcılara, 2009 yılındaki H1N1 grip salgını sırasındaki risklere ilişkin sorular verildi (Lee, Schwarz, Taubman & Hou, 2010).

ŞEKİL 6.4 Deneyler, çocukların mahkemede tanıklık yapmalarında olduğu gibi, gerçek hayattaki durumlara nasıl benzeyebilir?



İlk deney bir üniversite kampüsünde, ikincisi ise büyük alışveriş merkezlerindeki mağazalarda ve şehir merkezine yakın iş yerlerinde yapıldı. Katılımcılar, araştırma öncesinde uygulanan kısa bir anketi cevapladılar; sonra, kısa bir anketi cevaplamadan önce, bazıları seçkisiz olarak ya öksüren ve aksıranların (bunlar gerçekte deneycinin yardımcılarıydı) olduğu deneysel koşula ya da kontrol koşuluna (öksürüğü ve aksırığı olmayanlar koşuluna) atandılar. Sonuçlar bu basit değişimlemenin katılımcıların risk algısını etkilediğini gösterdi. Aksırmanın olmadığı koşula kıyasla, aksırmanın olduğu koşuldaki katılımcılar verilen ankette, 50 yaşından önce kalp krizi geçirme gibi ciddi bir hastalığa yakalanma, kaza geçirme ya da cinayet sonucu ölme risklerinin olduğunu daha fazla bildirdiler. İlginç şekilde, kontrol koşulundaki katılımcılar ile kıyaslandığında, öksürme koşulundaki katılımcılar, devletin "çevreci" işlere yatırım yapmasından daha çok, grip aşılarına harcama yapmasını desteklediler. Bu deney doğal bir ortamda yapıldığı için, "gerçek dünya" koşullarını temsil etmesi daha olasıdır. Bu nedenle, gerçek ortamlara genelleyebileceğimiz sonuçları, yapay ortamda oluşturulmuş deney sonuçlarından daha güvenilir buluruz.

Deneysel bulguların dış geçerliği aynı zamanda *kısmî tekrar* çalışması (kısmî replikasyon) ile de sağlanabilir. Kısmî tekrar çalışmaları çoğunlukla, bir olayın güvenilir bir şekilde ortaya çıktığı koşulları araştırma sürecinin rutin bir parçası olarak yapılır. Böyle bir tekrar, orijinal deneyin yönteminden biraz farklı bir yöntem uygulandığı zaman, orijinal deneyin sonuçlarına benzer sonuçlar elde edileceğini göstererek, dış geçerliğin sağlanmasına yardımcı olabilir. Aynı basit deneyin hem büyük bir kent üniversite hastanesinde hem de taşradaki küçük bir klinikte yapıldığını düşünelim. Katılımcılar ve deneylerin yapıldığı ortamlar çok farklıdır. Eğer bu farklı katılımcılar ve deney ortamı ile de aynı sonuçlar elde edilirse, sonuçların bu iki topluluğa ve ortama genellenebileceğini söyleyebiliriz. Deneyde sadece dış geçerlik olduğuna dikkat edin; bunlar, dış geçerliği olan *her iki* deneyde de ortaya çıkan *bulgulardır.*

Araştırmacılar aynı zamanda, *kavramsal tekrar* (replikasyon) yaparak da bulgularının dış geçerliğini sağlayabilirler. Bir çalışmadan genellemek istediğimiz şey, değişkenler arasındaki kavramsal ilişkilerdir; değişimlemeler, ortamlar ya da örneklemler gibi spesifik koşullar değildir (bk. Banaji & Crowder, 1989; Mook, 1983). Anderson ve Bushman (1997) kavramsal tekrar çalışmasının mantığını gösteren bir örnek veriyor. Beş yaşındaki çocuklarla, aşağılayıcı bir ifadenin ("mankafa!"), öfke ve saldırganlığı tetikleyip tetiklemediğini belirlemek için bir araştırma yapıldığını düşünün. Aynı aşağılayıcı ifadenin, 35 yaşındaki yetişkinlerde de aynı sonuca yol açıp açmadığını görmek için bir tekrar çalışması yapılabilir. Anderson ve Bushman'ın belirttiği gibi, 35 yaş grubunun sonuçları 5 yaş grubununkiyle aynı olmaz, çünkü bu ifade bu iki grup üzerinde aynı etkiyi yapmaz (s. 21). Yine de, eğer biz "sövmenin saldırgan davranışı arttırdığı" düşüncesinin dış geçerliğini sağlamak istersek, her evren için (5 ve 35 yaş grupları için) aynı anlama gelecek farklı ifadeler kullanabiliriz.

Anderson ve Bushman (1997) kavramsal düzeyde saldırganlıkla ilişkili değişkenleri incelediğinde, laboratuvar ortamında yürütülen deneylerden elde edilen bulgular ile gerçek hayattaki korelasyonel çalışmalardan elde edilen bulguların çok benzer olduğunu buldular. Yapay laboratuvar deneylerinin kesinlikle saldırganlık hakkında anlamlı bilgi sağladığı sonucuna vardılar, çünkü gerçek dünyadaki saldırganlıkta gözlenen aynı kavramsal ilişkileri gösterdiler. Dahası, laboratuvar deneyleri araştırmacılara, saldırganlığın potansiyel nedenlerini izole etme ve saldırganlığın ortaya çıkacağı veya çıkmayacağı zamana ait koşulları araştırma olanağı sağlar.

Gerçek dünyada ve laboratuvarda elde edilen sonuçlar birbiriyle çelişirse ne olur? Anderson ve Bushman'a (1997) göre bu çelişkilerin herhangi bir yöntemin zayıf olduğuna ilişkin bir kanıt olarak düşünülmemesi, aksine, saldırganlığa ilişkin teorilerimizi sadeleştirmede yardımcı olarak kullanılması gerekir. Bu çelişkili sonuçlar, bize her ortamda farklı psikolojik süreçlerin iş başında olabileceğini hatırlatması açısından önemlidir. Bu uyumsuz bulgulara ait anlayışımızı arttırırsak, saldırganlığa ilişkin anlayışımız da gelişecektir.

Kavramsal ya da kısmî tekrar çalışmaları yaparak psikolojideki *her* bulgunun dış geçerliğini sağlamak, gerçekte imkansızdır. Örneğin bir grup üniversite öğrencisi ile elde edilmiş bir deneysel bulguyu, daha yaşlı yetişkinlere, uzman kişilere, daha az eğitimli bireylere ve bunun gibi durumlara nasıl genelleriz? Bir olası yaklaşım, üzerinde düşünmeye değer görünüyor: Aksini düşünmek için

bir neden olmadıkça, davranışın zaman, denekler ve ortam açısından göreli olarak devamlı (benzer) olduğunu varsaymalıyız (Underwood & Shaughnessy, 1975). En sonunda, araştırma bulgularının dış geçerliğinin, kesin görgül kanıttan çok, bilim camiasının sağduyusu ile belirlenmesi muhtemeldir.

EŞLEŞTİRİLMİŞ GRUPLAR DESENİ

- Seçkisiz atama yapılan ve denek sayısının çok az olduğu durumda, etkili bir çalışma yapmak istiyorsak, karşılaştırılabilir gruplar oluşturmak için eşleştirilmiş gruplar deseni kullanılabilir.
- Denekleri bağımlı değişken açısından eşleştirmek (bir ön test gibi), eşleştirilmiş gruplar oluşturmada en iyi yoldur. Fakat eşleştirilen değişkenin puanları, bağımlı değişkenle korelasyonlu olmalıdır.
- Denekler bu şekilde eşleştirildikten sonra, seçkisiz olarak bağımsız değişkenin farklı koşullarına atanmalıdır.

Etkin bir araştırma, seçkisiz örneklemler deseninin, denekler arasındaki bireysel farkların seçkisiz atama yolu ile dengelenmesini garantileyen yeterli büyüklükte örneklemler gerektirir. Yani, seçkisiz gruplar deseninin varsayımı şudur: Bireysel farklar, gruplar arasında "ortalama" bir dağılım gösterir. Fakat bu ortalamaya doğru bir şekilde ulaşabilmek için ne kadar deneğe ihtiyaç var? Cevap: "Duruma göre değişir". Eğer örneklem homojen değil de heterojen bir topluluktan alınan deneklerden oluşturulmuşsa, bireysel farkları ortalama sayıya ulaştırmak için daha fazla denek gerekecektir.

Heterojen bir topluluktan alınan az sayıda denek test edildiğinde, seçkisiz atamanın denekler arasındaki bireysel farkları dengelemede çok etkin olmayacağından, büyük ölçüde emin olabiliriz. Yine de bu durum, psikolojinin bir çok alanındaki araştırmacıların karşılaştığı bir durumdur. Örneğin, bazı gelişim psikologları yeni doğmuş bebekler üzerinde araştırma yapar; başka bazıları ise yaşlıları araştırır. Hem yeni doğanlar hem de yaşlılar kesinlikle farklı evrenleri temsil ederler ve gelişim psikologları çoğunlukla, çalışabilecekleri sınırlı sayıda katılımcıya ulaşabilirler.

Bir alternatif de bu durumdaki araştırmacıların tekrarlı ölçüm kullanarak, tüm denekleri tüm deney koşullarına dağıtmasıdır (Bölüm 7'de tartışılacaktır). Yine de bazı bağımsız değişkenler her düzey için ayrı grupları gerektirir. Örneğin, varsayın ki araştırmacılar erken doğmuş bebekler için iki tür bakımı karşılaştırmak istiyorlar. Her bebeğe bu iki bakım türünün de uygulanması olanaksızdır. Bu ve benzeri bir çok durumda, araştırmacılar deneydeki ayrı grupları test etme ihtiyacı duyacaklardır.

Temel Kavram

Ne seçkisiz gruplar deseni ne de tekrarlı ölçümler deseni etkin şekilde kullanılamıyorsa, **eşleştirilmiş gruplar deseni** iyi bir alternatiftir. Eşleştirilmiş gruplar deseninin mantığı basit ve etkilidir. Denk gruplar oluşturmak için seçkisiz atamaya bel bağlamaktansa, araştırmacı denekleri eşleştirerek grupları denk hale getirir. Eşleştirmeye dayalı denk gruplar oluşturulmuşsa, eşleştirilmiş gruplar deseninin mantığı, seçkisiz gruplar desenindeki ile aynıdır (bk. Şekil 6.5). Eşleştirilmiş gruplar deseni kullanılırken, denekleri eşleştirmek için çoğu

ŞEKİL 6.5 Heterojen bir topluluktan alınmış az sayıda denek test edileceği zaman, seçkisiz atamanın denekler arasındaki farkları dengelemede etkili olması mümkün değildir (örneğin, yeni doğanlar). Bu durumda araştırmacılar, eşlenmiş gruplar desenini kullanabilirler.



zaman, bir eşleştirme değişkeni (ölçütü) ön test olarak kullanılır. Buradaki zorluk şudur ki grupları deneydeki sonuç ile ilişkili bir boyut açısından dengeleyecek bir ön test değişkeni bulunmalıdır. Bu nedenle, *eşleştirilmiş gruplar deseni, sadece iyi bir eşleştirme değişkeni* (ölçütü) *mevcut olduğu zaman faydalıdır*

En fazla tercih edilen eşleştirme değişkeni, deneyin kendisi içinde kullanılacak olan değişkendir. Örneğin, eğer deneydeki bağımlı değişken kan basıncı ise, katılımcılar deney başlamadan önce kan basıncı açısından eşleştirilmelidir. Eşleştirme, tüm katılımcıların kan basınçları ölçülerek, daha sonra aynı ya da benzer kan basıncı olan kişilerden, (deneydeki koşul sayısına bağlı olarak) ikili, üçlü, ya da dörtlü gruplar oluşturularak yapılır. Böylece, farklı gruplardaki katılımcılar, deneyin başlangıcında *ortalama olarak* birbirine denk kan basıncına sahip olacaktır. Bu durumda araştırmacılar, çalışmanın sonunda herhangi bir grupta gözlenen kan basıncı farkını, ikna edici şekilde, uygulanan tedaviye bağlayabilecektir. (tabi diğer potansiyel değişkenlerin sabit tutulmuş olduğunu ya da dengelenmiş olduğunu varsaymak koşulu ile).

Bazı deneylerde, temel bağımlı değişken denekleri eşleştirmek için kullanılamayabilir. Örneğin, katılımcılara bir bulmacayı çözmeleri için farklı yaklaşımların öğretildiği bir deney düşünün. Eğer bireylerin bu bulmacayı çözmelerinin ne kadar zaman aldığını görmek için bir ön test verildiyse, katılımcılar muhtemelen ön test sırasında bulmacayı çözmeyi öğrenecektir. Bu durumda, değişimlemenin ardından, farklı grupların bulmacayı çözme hızlarındaki farkı gözlemlemek olanaksız olacaktır. Bu nedenle eşleştirme değişkeni olarak bir

sonraki en iyi alternatif, aynı *sınıf ya da kategoriden* bir ölçüt kullanmaktır. Problem çözme örneğimizde, katılımcılara deneyde kullanılacak olan bulmacadan farklı bir bulmaca çözdürüldüğü durumda, katılımcılar bu performansları açısından eşleştirilebilirler. Eşleştirmede daha az tercih edilen, fakat hâlâ olasılık dahilinde olan bir alternatif de bağımlı değişkenden *farklı sınıftan* bir ölçüt kullanmaktır. Bizim problem çözme deneyimiz için, katılımcılar mekânsal yetenek testi gibi bir genel yetenek testi açısından eşleştirilebilirler. Bu alternatifleri kullanırken, yine de araştırmacılar eşleştirme değişkeninin performansının, deneyde kullanılan bağımlı değişkenin performansı ile ilişkili olduğunu doğrulamalıdır. Genelde eşleştirme değişkeni ve bağımlı değişken arasındaki ilişki azaldıkça, eşleştirilmiş gruplar deseninin avantajı, seçkisiz gruplar desenine kıyasla, azalır.

İyi bir eşleştirme değişkeninin olduğu bir durumda bile, eşleştirme tekniği, denk gruplar oluşturmak için yeterli olmayabilir. Örneğin, prematüre bebeklerin vücut ağırlıklarını arttırmak amacıyla kullanılan iki farklı bakım yöntemini karşılaştırmak için eşleştirilmiş gruplar deseninin kullanıldığını düşünün. Altı çift prematüre bebek başlangıçtaki vücut ağırlıkları açısından eşleştirilebilir. Bu durumda yine de, eşleştirme değişkeninin (ilk ağırlık) dışında, bebeklerin ağırlıklarını etkileyebilecek başka özellikleri kalabilir. Örneğin, iki prematüre bebek grubu genel sağlık ya da ebeveynlerine bağlanma derecesi açısından denk olmayabilir. Bu nedenle, eşleştirilmiş gruplar deseninde eşleştirilen değişken dışında diğer potansiyel etkenleri dengeleyebilmek için seçkisiz atama yöntemini de kullanmak önemlidir. Spesifik olarak, bebekler vücut ağırlıkları açısından eşleştirildikten sonra, her bebek çiftindeki bireyler iki tedavi grubundan birine seçkisiz atanmış olacaktır. Sonuç olarak, *her koşul için ayrı gruplar oluşturulmasını gerektiren, ancak denek sayısının az olduğu bir deneyde, eğer uygun bir eşleştirme değişkeni varsa, eşleştirilmiş gruplar deseni seçkisiz gruplar deseninden daha iyi bir seçenektir.*

DOĞAL GRUPLAR DESENİ

- Doğal gruplar deseni oluşturulurken, bireysel farklar değişkenleri (ya da denek değişkenleri) değişimlenmez, seçilir.
- Doğal gruplar deseni, araştırmacıların doğal gruplar değişkenleri ve bağımlı değişkenler arasındaki birlikte değişimleri (covariation) aradığı, bir tür korelasyonel araştırmadır.
- Doğal gruplar deseni etkilerine ilişkin nedensel çıkarımlar yapılamaz, çünkü grup farklılıklarının mümkün alternatif açıklamaları vardır.

Temel Kavram

Psikolojinin birçok alanında araştırmacılar **bireysel farklar değişkenleri** (individual differences variables) adı verilen bağımsız değişkenlerle ya da *denek değişkenleri* ile ilgilenirler. Bir bireysel farklar değişkeni, bireyden bireye değişen bir kişisel özelliktir. Dinî eğilim, bireysel farklar değişkenine bir örnektir. Araştırmacılar bu değişkeni katılımcıları Katolik, Musevi, Protestan ya da diğer gruplara seçkisiz yöntemle atayarak değişimleyemezler. Bunun yerine, araştırmacılar dinî eğilim değişkenini, doğal olarak bu gruplara ait bireyleri sistemli

şekilde seçerek " kontrol" ederler. Cinsiyet, içedönüklük-dışadönüklük, ırk ya da yaş gibi bireysel farklar değişkenleri, psikolojinin bir çok alanında kullanılan önemli bağımsız değişkenlerdir.

Düzeyleri *değişimlenen* bağımsız değişkenleri kapsayan deneylerden, düzeyleri *seçilen* bağımsız değişkenleri kapsayan deneyleri ayırt etmek önemlidir.

Temel Kavram

Bireysel farklar değişkenleri gibi düzeyleri seçilmiş olan bağımsız değişkenleri kapsayan deneylere, **doğal gruplar deseni** adı verilir. Doğal gruplar deseni, sıklıkla bağımsız değişkeni doğrudan değişimlememizi önleyen etik ve pratik kısıtlamaların olduğu durumlarda kullanılır. Örneğin, önemli bir ameliyatın, o ameliyattan sonra ortaya çıkan bir depresyon üzerindeki etkileri ile ne kadar ilgilenirsek ilgilenelim, seçkisiz atanmış bir psikolojiye giriş öğrencileri grubuna önemli bir ameliyat yapmamız, sonra da bunların depresyon belirtilerini ameliyat olmamış bir başka grubunkilerle karşılaştırmamız etik olmaz. Benzer şekilde, boşanma ile duygusal sorunlar arasındaki ilişki ile ilgileniyorsak, bu konuyu araştırırken, bazı insanları ayrılmaları için seçkisiz olarak atayamazdık. Ama doğal gruplar desenini kullanarak, yine de cerrahi operasyon geçirenleri geçirmeyenlerle karşılaştırabiliriz. Aynı şekilde, ayrılmayı seçmiş olanlar, evli kalmayı seçmiş olanlarla karşılaştırılabilir.

Araştırmacılar bilimsel yöntemin ilk iki amacını gerçekleştirmek üzere doğal gruplar desenini kullanırlar: Betimleme (description) ve kestirim (prediction). Örneğin, araştırmalar göstermiştir ki boşanmış insanların psikiyatrik bakım almaları olasılığı evli, dul ya da bekâr olanlardan daha fazladır. Bu gibi araştırma sonuçlarına dayanarak, ayrılmış ya da boşanmış bireyleri, duygusal bozukluklar açısından betimleyebiliriz ve hangi grubun daha fazla duygusal bozukluk yaşayabileceğini kestirebiliriz.

Doğal gruplar deseninin sonuçları nedensel açıklamalar yapmak için kullanılırsa, ciddi problemler ortaya çıkabilir. Örneğin, ayrılmış insanların evli insanlara kıyasla daha fazla psikiyatrik bakım almasının daha olası olduğu bulgusu, bu iki etkenin birlikte değiştiğini gösterir. Bu bulgu, boşanmanın duygusal bozukluklara yol açabildiği anlamına gelebilir. Fakat boşanmanın duygusal bozukluklara yol açtığı sonucuna varmadan önce, nedensel bir çıkarım yapabilmek için, nedensel ilişkinin zorunlu bir koşulu olan zaman-sıra koşulunun karşılandığından emin olmamız gerekir. Boşanma mı duygusal bozukluktan önce geliyor, yoksa duygusal bozukluk mu boşanmadan önce geliyor? Bir doğal gruplar deseni bu sorunun cevabını vermez.

Nedenselliğin üçüncü koşulu olan muhtemel alternatif açıklamaları elemek istediğimiz zaman da doğal gruplar deseni sorun çıkarır: Doğal gruplar deseniyle araştırılan bireysel farklar çoğunlukla karıştırıcıdır - katılımcıların oluşturduğu grupların, onları sınıflandırmak için kullanılan değişken *dışında* bir çok açıdan farklı olmaları olasılığı vardır. Örneğin, boşanan bireyler ve evli kalan bireyler medenî durumları dışında dinî adetleri ya da ekonomik koşulları gibi bir çok özellik açısından farklı olabilirler. Bu nedenle, boşanmış bireyler ile evli bireyler arasında gözlenen herhangi bir fark boşanmaya değil, bu diğer özelliklere bağlı olabilir. "Doğa" eli ile yapılan bu değişimleme nadiren, iç geçerliği sağlamak için kontrol ettiğimiz değişkenler gibi, değişimlenebilir.

Doğal gruplar deseninde nedensel çıkarımlar yapmaya yönelik yaklaşımlar vardır. Etkin bir yaklaşım, bireysel farkların bağımsız değişkenlerle birlikte araştırılmasını gerektirir. Bir deneyde birden çok bağımsız değişkenin kombinasyonu, Bölüm 8'de betimlenecek olan karmaşık bir desenin kullanımını gerektirir. Şimdilik, doğal gruplar desenine dayanan nedensel çıkarımlarda bulunmak tehlikeli bir girişim olabilir. Bu tür desenler bazen "deney" olarak ifade edilse de, bireysel farklar değişkenini içeren bir deney ile değişimlenmiş bir değişkeni içeren bir deney arasında önemli farklar vardır.

ÖZET

Araştırmacılar teorilerden çıkardıkları hipotezleri test etmek için deneyler yaparlar ama deneyler aynı zamanda uygulamalı alanlardaki uygulamaların ve programların etkinliğini test etmek için de kullanılır. Kontrol teknikleri, koşulların sabit tutulması ve dengeleme uygun şekilde uygulanırsa, deneysel yöntem sebep-sonuç ilişkilerini belirlemek için idealdir.

Bu bölümde bu kontrol tekniklerinin, bağımsız değişkenin farklı düzeylerini temsil eden farklı uygulamaların, farklı gruplara verildiği deneylerdeki uygulamasına odaklandık (bk. Şekil 6.6). Seçkisiz gruplar deseninde gruplar, seçkisizleştirme teknikleri kullanılarak oluşturulur; böylece, deneyin başlangıcında denk gruplar elde edilir. Eğer değişimlemden sonra gruplar farklı bir performans gösterirlerse ve diğer tüm koşullar sabit tutulduysa, bu farkın nedeninin bağımsız değişken olduğu kabul edilebilir. Seçkisiz atama, denk gruplar oluşturmada en yaygın kullanılan yöntemdir. Yine bu yöntem, deneydeki koşullar arasında, bireysel farkların dengelenmesini sağlama girişimidir. Seçkisiz atama yapmak için en yaygın teknik blok seçkisizleştirmedir.

Bağımsız grupları test eden deneylerin iç geçerliğine yönelik çeşitli tehditler vardır. Grupları bütün hâlinde test etmekten, bu gruplar seçkisiz atanmış olsalar bile kaçınmak gerekir; çünkü bu grupların karıştırıcı etkiye neden

ŞEKİL 6.6 Bu bölümde üç bağımsız değişken gruplar desenini tanıttık.



olmaları olasılığı yüksektir. Farklı uygulama odaları ya da farklı araştırmacılar gibi dış değişkenlerin etkisinin, bağımsız değişkenin etkisiyle karışması engellenmelidir.

Seçkisiz gruplar deseninde iç geçerliğe daha ciddi bir tehdit, denekler araştırmayı tamamlamakta başarısız olduklarında ortaya çıkar. Eğer denekler farklı koşullarda farklı nedenlerle kaybediliyorsa ve denek kaybının nedeni, deneklerin deney sonucu ile ilişkili bazı özellikleri ise, bu durumda seçici denek kaybı ortaya çıkar. Önlem olarak, sadece deneyi tamamlama olasılığı olan denekleri seçebiliriz veya kayıp yaşanmayan gruptan kaybedilenlere denk denekleri gruptan çıkararak bu durumu telafi edebiliriz. İşlemler sırasında ortaya çıkan bir hatanın denek kaybına yol açması ile ortaya çıkan mekanik denek kaybı daha az sorun yaratan bir durumdur. Talep özellikleri ve araştırmacı etkileri uygun deneysel işlemler etkili şekilde kullanılarak azaltılabilir, fakat kontrolü en etkili şekilde sağlamanın teknikleri, plasebo kontrol ve çift kör işlemleridir.

Veri analizi ve istatistik, bir bağımsız değişkenin davranış üzerinde etkileri olduğunu öne sürmek için tek bir deneyin sonuçlarının kullanılıp kullanılamayacağını belirlemek amacıyla yapılan tekrar çalışmasına (replication) alternatiftir. Veri analizi hem betimleyici hem de çıkarımsal istatistik kullanımını içerir. Bir deneyin sonuçlarını betimlemek için, tipik olarak ortalama, standart sapma ve etki büyüklüğü ölçümleri kullanılır. Meta-analiz önemli bir araştırma sorusu hakkında çok sayıda deneyden elde edilen sonuçların nicel bir özetini elde etmek için etki büyüklüğü ölçümünü kullanır.

Çıkarımsal istatistikler veri analizinde önemlidir, çünkü araştırmacıların bir deneyde elde edilen farkların şans eseri mi, yoksa bağımsız değişkenin etkisine bağlı olarak mı ortaya çıktığına karar vermeye yarar. Güven aralığı ve sıfır hipotezi testi, araştırmacıların deneylerde analiz yapmak için kullandıkları iki etkin istatistik tekniğidir. Ama unutulmamalıdır ki istatistiksel analiz, deneysel bulguların anlamlı veya pratik açıdan önemli olacağını garanti etmez. Tekrar çalışması, bir araştırma bulgusunun güvenirliğini test etmek için en güvenilir yol olarak kalmaya devam eder.

Araştırmacılar aynı zamanda deneysel bulguların dış geçerliğini sağlamak için de uğraşırlar. Psikoloji teorilerini test ederken araştırmacılar, iç geçerliğe, dış geçerlikten daha fazla önem verirler. Bulguların dış geçerliğini sağlamak için araştırmacıların kullanması gereken etkin bir yaklaşım, bulgularını genellemek istedikleri tüm özellikleri temsil eden bir örneklem seçmektir. Alan çalışmaları yaparken araştırmacılar, araştırma bulgularının gerçek dünyaya genelleme olasılığını arttırabilirler. Dış geçerlik sağlamak için kullanılan iki yaygın yol, kısmî ve kavramsal tekrar çalışmalarıdır.

Eşleştirilmiş gruplar deseni, ancak şu hâllerde seçkisiz gruplar desenine bir alternatiftir: Denek sayısının az olduğu, iyi bir eşleştirme değişkeninin olduğu, ve her deney koşulu için ayrı gruplar gerektiği zaman. Eşleştirilmiş gruplar desenine ilişkin en büyük problem, grupların sadece eşleştirme değişkeninin ölçtüğü özellik açısından dengelenmesidir. Doğal gruplar deseninde, araştırmacılar bağımsız değişkenin düzeylerini (çoğunlukla bireysel farklar ya da denek değişkenleri) seçerler ve bağımsız değişken ile davranışın diğer yönleri arasındaki

sistematik ilişkileri araştırırlar. Zorunlu olarak, doğal gruplar deseni deneklerin özellikleri ve performansları arasındaki korelayonu araştırır. Bu tür korelasyonel araştırma desenleri, nedensel çıkarımların yapılmasına uygun değildir.

TEMEL KAVRAMLAR

iç geçerlik 181
bağımsız gruplar deseni 182
seçkisiz atama 182
seçkisiz gruplar deseni 182
blok seçkisizleştirme 188
iç geçerliği tehdit eden etkenler 189
mekanik denek kaybı 191
seçici denek kaybı 191
çift kör işlemi 194
deneyci etkileri 194
plasebo kontrol grubu 194
tekrar çalışması 196
etki büyüklüğü 198
Cohen d 199
meta-analiz 199
istatistiksel anlamlılık 202
sıfır hipotezi anlamlılık testi (SHAT) 202
güven aralığı 203
eşleştirilmiş gruplar deseni 209
bireysel farklar değişkeni 211
doğal gruplar deseni 212

GÖZDEN GEÇİRME SORULARI

1 Psikologların niçin araştırma yaptıklarını açıklayan iki neden belirtin.
2 Kontrol tekniklerinden değişimleme, koşulları sabit tutma ve dengeleme, nedensel bir çıkarım için zorunlu olan üç koşulun yerine getirilmesine nasıl katkıda bulunur?
3 Denk grupların, seçkisiz gruplar deseninin bu kadar temel bir özelliği olmasının nedenini açıklayın ve araştırmacıların denk grupları nasıl oluşturduklarını betimleyin.
4 Blok seçkisizleştirmede "blok" neye karşılık gelir ve bu işlem bize ne sağlar?
5 Seçici denek kaybının deneyinizde problem oluşturacağını fark ettiğiniz zaman, hangi önlemleri alırsınız?
6 Plasebo kontrol ve çift kör işlemlerinin, talep özelliklerini ve deneyci etkilerini kontrol etmede nasıl kullanılabileceğini açıklayınız.
7 Meta-analiz neden araştırmacıların psikolojinin ilkelerine ilişkin daha güçlü sonuçlar çıkarmasına olanak sağlar?
8 Çıkarımsal istatistik kullanılarak anlamlı bulunan bir sonucun, bir deneydeki bağımsız değişkenin etkisi hakkında ne söylediğini açıklayınız.
9 Bir deneydeki iki koşula ilişkin ortalamalar arasındaki farkı test ederken, eğer güven aralıkları örtüşmediyse, bu durumdan nasıl bir sonuç çıkaracağınızı açıklayın.
10 Bir araştırma bulgusunun dış geçerliğinin olması için araştırmacının izleyeceği dört yolu kısaca açıklayın.
11 Kısaca eşleştirilmiş gruplar deseninin mantığını açıklayın ve eşleştirilmiş gruplar deseninin seçkisiz gruplar desenine kıyasla daha iyi bir alternatif olduğu üç koşulu belirtin.
12 Bireysel farklar değişkeni, değişimlenen bağımsız değişkenlerden nasıl farklılaşır ve bu fark, doğal gruplar deseninde nedensel çıkarımlar yapmayı nasıl zorlaştırır?

DÜŞÜNDÜRÜCÜ SORULAR

1. Bir deneyci seçkisiz gruplar deseninin kullanıldığı bir deneyde, sunulan uyarıcı hızının, insanların uyarıcıyı tanıma yeteneği üzerindeki etkisini araştırmayı planlıyor. Bağımsız değişken sunum hızıdır ve dört düzeyde değişimlenecektir: Çok hızlı, hızlı, yavaş ve çok yavaş. Deneyci sizden yardım istiyor ve deney hakkında aşağıdaki önerilerde bulunuyor:
   - **A** Deneyci dört koşulun her birinde dört katılımcı olacak şekilde bir blok seçkisizleştirme programı hazırlamanızı istiyor. Bunu yapabilmek için, ekte (Tablo A.1) yer alan seçkisiz sayı tablosundan alınmış seçkisiz sayıları kullanabilirsiniz.

     1-5-6-6-4-1-0-4-9-3-2-0-4-9-2-3-8-3-9-1
     9-1-1-3-2-2-1-9-9-9-5-9-5-1-6-8-1-6-5-2
     2-7-1-9-5-4-8-2-2-3-4-6-7-5-1-2-2-9-2-3

   - **B** Deneyci, kendilerine çok hızlı sunum hızı ile sunulan görevi başarılı bir şekilde yapabileceğinden emin olmak için araştırmasını, verdiği reaksiyon zamanı testinden geçenlerle sınırlandırmayı düşünüyor. Sanki sadece bu sınırlandırılmış katılımcılar kümesi test ediliyormuş gibi düşünerek, karşılaşılabileceği risklerin neler olabileceğini belirleme aşamasında, araştırmacının hangi etkenler üzerinde düşünmesi gerektiğini açıklayınız.
   - **C** Deneyci katılımcıları iki ayrı odada test etmesi gerektiğini fark ediyor. Deneyci bu farklı oda koşullarında test etme işlemini, bu dış değişkenin karıştırıcı etkilerini önlemek için nasıl düzenlemelidir?

2. Bir deneyci kadınların kendi beden imgelerine ilişkin duygularının, eşik altı beden imgesi sunumlarından etkilenip etkilenmediğini belirlemek üzere araştırma yaptı. Bir imgenin 15 milisaniyede (ms) sunumu, insanların imgenin farkında olmaları için çok hızlıdır; bu da uyarıcının sunumunun eşik altı ya da bilinçli farkındalığın dışında olması anlamına gelir.

   Kadınlardan (N= 60), ekrandaki kelimeler hakkında karar verme hızının ölçüldüğü bir deneye katılmaları istendi. "Nötr kelimeler (beden imgesi ile ilişkili olmayan), ekranın merkezinde 3 saniye süre ile sunuldu ve katılımcılar, eğer kelime büyük harfle yazılmışsa, soldaki tuşa ve küçük harfle yazılmışsa sağdaki tuşa bastılar. Tüm katılımcılar, seçkisiz olarak yarısı büyük harfle ve yarısı da küçük harfle yazılmış olan aynı kelimeleri gördüler. Yine de denemelerin seçkisiz olarak yarısında, kelime görünmeden 15 ms önce bir görüntü sunuldu. Bu görüntüler zayıf (Z) veya aşırı kilolu (AK) kadınların görüntüleri ya da bir kitap veya bitki gibi nötr (N) görüntülerdi; kadınlar bu koşullara seçkisiz olarak atandılar. Her gruptaki kadın, bu görüntülerin sadece bir tipi ile karşılaştı. 20 katılımcının olduğu her koşulda, her bir kadın 50 görüntü gördü.

   Bilgisayar görevini tamamladıktan sonra, katılımcılar beden memnuniyetsizliği ölçümü ile ilgili birkaç anket cevapladılar. Kadınlardan, her birisinde çok zayıftan çok şişmana doğru dizilmiş yedi kadın şeklinin bulunduğu bir sırayı seçmeleri, ve bu kadın şekilleri arasından ikisini işaretlemeleri istendi: İlki, kendilerinin görünmek istedikleri ideal şekil, ikincisi ise gerçekte kendi bedenlerini nasıl algıladıklarını en iyi ifade eden şekil olacaktı. Beden memnuniyetsizliğinin işevuruk tanımı olarak, kadınların işaretledikleri ideal ve gerçek şekil arasındaki şekil sayısı belirlendi (0, memnuniyetsizlik olmadığını, negatif puanlar daha zayıf olma arzusunu ve pozitif puanlar ise kilo alma isteğini gösterir).

   Zayıf beden imgesi grubuna ilişkin beden memnuniyetsizliği ortalaması –1.25 (SD= 1.07), aşırı kilolu imge grubuna ilişkin puan ortalaması –0.75 (SD=1.16) ve nötr imge grubuna ilişkin puan ortalaması –0.20 (SD= .83) oldu.
   - **A** Bağımsız ve bağımlı değişkenleri belirtecek şekilde, bu araştırmanın desenini betimleyin. Araştırmacının katılımcılardan niçin aynı nötr kelimeleri her koşulda değerlendirmelerini istediğini açıklayın ve deneycinin koşullara seçkisiz atayarak kontrol ettiği faktörlere örnekler verin.
   - **B** Her koşuldaki ortalamaları kullanarak bağımsız değişkenin etkisini nasıl belirlersiniz? Standart sapmalar size deneydeki beden memnuniyetsizliği değerlendirmelerine ilişkin ne söylemektedir?
   - **C** F testinde imge tipi etkisine ilişkin p değeri, p=.009'dur. Bu olasılık, bağımsız değişkenin etkisi hakkında ne anlama gelir?
   - **D** Zayıf imge koşulundaki .95 güven aralığı (GA) –1.71 ile – 0.79; aşırı kilolu imge koşulunda GA –1.21 ile –0.29 ve nötr imge koşulunda GA –0.66 ile 0.26'dır. Bu güven aralıklarını karşılaştırarak, deneydeki üç grubun evren ortalamaları hakkında hangi tahminlerde bulunabilirsiniz?
   - **E** Zayıf ve nötr gruplara ilişkin ortalamalar arasındaki farkın etki büyüklüğü d=.85'tir.

Testin istatistiksel anlamlılığı ve güven aralıklarının karşılaştırılması ile öğrendiklerimizin ötesinde bu etki büyüklüğü, nötr imgelere kıyasla, eşik altı zayıf imgelere maruz kalmanın etkisine ilişkin ne söyler?

3 Prematüre yeni doğanlar vücut ısılarını düzenlemekte, düşük doğum ağırlıkları ve düşük yağları nedeniyle, sıklıkla sorun yaşarlar. Araştırmacılar ılık suya daldırmanın, yaygın uygulama olan süngerle banyo yaptırmaya kıyasla, vücut ısısını daha iyi düzenleyeceği hipotezini kurdular. Sekiz yeni doğan hipotezi test etmek için uygundu; sekiz bebeğin her birisi, seçkisiz olarak yoğun bakım birimindeki bu iki koşuldan birisine atandı. Bağımlı değişken banyodan 15 dakika sonra ölçülen vücut ısısıydı. Hipotezin tersine, gruplar arasında, banyo sonrası vücut ısısında hiçbir fark bulunmadı.

A Bir hemşire, iki yeni doğan grubunun çok farklı olduğunu fark etti. Ortalamalar açısından bakıldığında, suya daldırma koşuluna seçkisiz atanan yeni doğanlar, süngerle banyo yaptırılan koşuldakilere kıyasla daha az kilodaydılar ve daha erken doğmuşlardı. Peki katılımcılar koşullara seçkisiz olarak atandığı halde, nasıl oldu da bu iki grup arasında böyle bir fark ortaya çıktı ve bu durumda bu fark deneyin sonuçlarını nasıl etkilemiş olabilir? Açıklayın.

B Araştırmacılar hipotezi tekrar test etmeye karar verdiler. Bu defa ayrıca, yeni doğanların ağırlığı üzerinde kafa yordular. Ağırlıkları (pound cinsinden) 4.6, 5.0, 4.1, 4.0, 5.4, 5.8, 4.4, 4.9, 6.2, 5.2, 4.2, 4.6. olan on iki yeni doğan, araştırma için uygun bulundu. Bu çalışmayı yürütmek için önereceğiniz araştırma desenini, ve yeni doğanların deney koşullarına nasıl atanması gerektiğini açıklayın.

C Varsayın ki bir sosyal hizmet görevlisi, ebeveynlik becerilerini öğrettiği ev gezileri sırasında banyo değişkenini test ediyor. Görevli, vücut ısısını evde korumanın, küvezlerin olduğu yoğun bakım birimlerinde korunmasından daha önemli olduğunu düşündü. 20 prematüre yeni doğan ailesi, bebeklerini belirli bir derecedeki suya daldırma koşuluna, 20 aile ise süngerle banyo yaptırma koşuluna seçkisiz olarak atandı. Sosyal hizmet görevlisi ailelere banyodan 15 dakika sonra bebeğin ısısını (bağımlı değişken) nasıl ölçeceğini de öğretti. Bu araştırmayı yürütürken, görevliyi böyle bir araştırma düzeni kurmaya yönelten amacın ne olabileceğini bulun. Eğer her iki deneyde de süngerle banyo yaptırmaya kıyasla, suya daldırma banyosu uygulanan yeni doğanların vücut ısısı daha yüksek bulunursa, böyle bir sonucu nasıl yorumlarsınız?

4 Diyabetik erkek hastalara kilo kaybettirmek için yapılan iki diyeti karşılaştırmak için bir çalışma yapıldı. Her diyet, hastanın aldığı günlük kalori miktarını 1800 kaloriye kadar düşürdü. Kontrol koşulunda, hastalar izin verilen yiyecek listesinden kendi yemeklerini seçtiler, ama yine de günlük kalori limitini aşmalarına izni verilmedi. Tedavi koşulunda ise insülin seviyesindeki dalgalanmaları en aza indirmek için, yemekler aynı listeden seçildi; bu amaçla diyet yemekleri, çeşitli karbonhidrat ve proteinlerden belirli oranlarda alınarak oluşturuldu. Her koşula kırk erkek katılımcı seçkisiz olarak atandı. Tüm katılımcıların insülin seviyeleri izlendi. Deney 6 ay sürdü; bu süre içinde kontrol koşulundaki 15 erkek, çeşitli nedenlerle araştırmadan ayrıldı (örneğin, insülin ile baş etme zorlukları, diyabete ilişkin komplikasyonlar gibi nedenlerle). Tedavi koşulundakilerin ise üçü araştırma dışı kaldı; ikisi şehir dışına gitti ve üçüncü kişi ise iş saatleri ile çakıştığı için randevularına gelemedi. Altı ayın sonunda, kontrol koşulunda çalışmayı tamamlayan 25 erkeğin ortalama kilo kaybı, 11 pound ve tedavinin yapıldığı koşulda çalışmayı tamamlayan 37 erkeğin ortalama kilo kaybı 10 pound idi. Araştırmacı farklı kaynaklardan farklı oranlarda kalori alımını spesifikleştiren bu uygulamanın (örneğin, kompleks karbonhidratlar, proteinler), diyabetik erkeklerde kilo kaybına yardımcı olmadığı sonucuna vardı.

A Bu deneyin iç geçerliğine yönelik olası bir tehdidi bulun ve bu tehdidin, araştırmanın sonuçlarını nasıl etkilemiş olabileceğini açıklayın.

B Çalışmadaki 80 erkek için ön test ölçümünün, gerektiğinde düzenlenebilen ve düzeltilebilen insülin seviyesi olduğunu varsayın. Bu ön test puanlarının, iç geçerliğe bir tehdit oluşturup oluşturmadığını nasıl belirlersiniz?

İleri Alıştırma I'in Cevapları

1 Bushman (2005) araştırmasında, bağımsız değişken olarak seçtiği televizyon programı türünü, dört düzeyde değişimledi: Şiddet içeren, cinsellik içeren, şiddet ve erotizm içeren ve nötr düzeyler.

2 Bushman (2005) birkaç faktörü sabit tuttu: Her koşulda aynı reklamları kullandı, katılımcılar aynı ortamlarda küçük gruplar hâlinde test edildi, ve reklamlar her programda yaklaşık aynı noktada verildi.

3 Bushman (2005) katılımcıları koşullara seçkisiz şekilde atayarak, katılımcıların özelliklerini dört koşulda dengeledi. Böylece her koşul ortalama olarak, bellek yeteneği ve katılımcıların televizyon programı ve ürünlerine maruz kalışı açısından denk oldu. Bushman aynı zamanda, TV programları sırasında reklamların gösterime girdiği yere bağlı olarak ortaya çıkabilecek olası etkileri dengelemek için iki seçkisiz reklam dizisi kullandı.

İleri Alıştırma II'nin Cevapları

1 Bu kitabın yazarlarından biri bu alıştırmayı tamamladığında, deneysel grup için ortalama 5.65 ve kontrol grubu için 5.35 değerlerini buldu. İki grup, bellek becerileri ortalaması açısından denkti (bu ortalamalar arasındaki farkın anlamlı olup olmadığını belirlemek için bir *t* testi yapılabilir).

2 Deneysel grupta 2 puanlı üç "katılımcı" vardı. Bunlar deney dışı kaldığında, bellek becerilerinin yeni ortalaması 6.4'tü. Kontrol grubunun ortalaması olan 5.35 ile karşılaştırıldığında, seçici denek kaybından sonra deneysel grubun ortalama bellek becerisi daha fazlaydı.

3 Üç deneğin kaybını telafi etmek için, benzer "katılımcılar" kontrol grubundan da çıkarıldı (2, 1 ve 1'in puanları). Kontrol grubunun yeni ortalaması 6.06 idi. Bu, iki grubun başlangıçtaki denklik düzeyini yükseltti

4 Bu kitabın yazarlarından birinin yaptığı düzenlemeye göre, dört grubun ortalamaları (1) 5.6, (2) 4.8, (3) 5.3 ve (4) 6.3 şeklindedir; bu, gruplar arasında bellek becerisi puanları ortalamaları açısından daha fazla değişkenlik olduğunu göstermektedir. Koşullara seçkisiz atanan katılımcı sayısı ne kadar az ise, seçkisiz atama ile denk grupların oluşturulması o kadar zordur. Şimdi bunları bir kenara bırakalım ve Bölüm 6'ya geri dönelim.

Düşündürücü Soruların Cevapları

A Birinci aşama, 1'den 4'e kadar sırası ile koşullara atamaktır: 1 = Çok hızlı; 2 = Hızlı; 3 = Yavaş; ve 4 = Çok yavaş. Bu durumda, seçkisiz sayıları kullanarak, 1'den 4'e kadar olan sayılardan oluşan dört dizi seçin. Bunu yaparken, sayıların dörtten büyük, dizide daha önce yer almış, ve her hangi bir sayının tekrarı olmamasına dikkat edin. Örneğin, seçtiğiniz ilk sayı 1 ise, 1'den 4'e kadar olan dizideki tüm sayılar seçilene kadar, 1'in tüm tekrarlarından kaçının. Böylece soldan sağa seçkisiz sayılar dizisi üzerinde çalışarak, seçkisiz bloklamanın dört bloğuna ilişkin ardışık dört dizi elde ettik. Her blok için koşulların sıralaması gösterilmiştir. Blok seçkisizleştirme yöntemi, deneydeki ilk 16 katılımcının koşullarının test edilme sırasını belirler.

Blok 1: 1-4-3-2 Çok Hızlı, Çok Yavaş, Yavaş, Hızlı

Blok 2: 4-2-3-1 Çok Yavaş, Hızlı, Yavaş, Çok Hızlı

Blok 3: 1-3-2-4 Çok Hızlı, Yavaş, Hızlı, Çok Yavaş

Blok 4: 2-3-4-1 Hızlı, Yavaş, Çok Yavaş, Çok Hızlı

B Araştırmacı seçici denek kaybını önlemek için önemli bir aşamadadır, fakat sadece katı tepki zamanı testinden geçen katılımcıları deneye alması, elde edilen bulguların dış geçerliği açısından risk oluşturur.

C Blok seçkisizleştirme planındaki her bir blok, her bir odaya atanarak, odalar dengelenebilir. Her bir odaya atanan blok sayısı daima eşittir, ama esas mesele bu değildir. Etkili bir dengeleme için, her bir odada çeşitli bloklar test edilmelidir.

BÖLÜM YEDİ

Tekrarlı Ölçümler Desenleri

BÖLÜMÜN ANA HATLARI

GENEL BAKIŞ
ARAŞTIRMACILAR NEDEN TEKRARLI ÖLÇÜMLER DESENLERİNİ KULLANIR?
TEKRARLI ÖLÇÜMLER DESENLERİNDE ALIŞTIRMA ETKİLERİNİN ROLÜ
Alıştırma Etkilerini Belirleme
Tamamlanmış Desende Alıştırma Etkilerini Dengeleme
Tamamlanmamış Desende Alıştırma Etkilerini Dengeleme
TEKRARLI ÖLÇÜMLER DESENLERİNDE VERİ ANALİZİ
Sonuçların Betimlenmesi
Verilerin Anlamını Onaylama
AYRIMSAL AKTARMA SORUNU
ÖZET

GENEL BAKIŞ

Şimdiye kadar deneklerin sadece bir deney koşulunda yer aldığı deneylerden bahsettik. Denekler seçkisiz gruplardaki bir koşula seçkisiz olarak atanırlar veya doğal gruplar deseninde olduğu gibi, gruplardan birisinin üyesi olarak seçilirler. Bu bağımsız gruplar desenleri, geniş bir aralıktaki bağımsız değişkenlerin etkilerini araştırmaya uygun güçlü araçlardır. Bununla birlikte bazen her bir deneğin bir deneyin tüm koşullarında yer almasının daha etkili olduğu durumlar vardır. Bu desenler **tekrarlı ölçümler desenleri** (repeated measures designs) adını alır (veya denek içi desenler). Bir bağımsız gruplar deseninde, ayrı bir grup kullanılması, deneysel uygulamanın verildiği grup için kontrol işlevi görür. Bir tekrarlı ölçümler deseninde ise denekler kendi *kendilerinin kontrol grubu işlevini görürler*, çünkü hem deneysel hem de kontrol koşullarında yer alırlar.

Temel Kavram

Bu bölüme araştırmacıların neden tekrarlı ölçümler desenlerini seçtiklerini araştırarak başlıyoruz. Daha sonra tekrarlı ölçümler desenlerinin temel özelliklerinden birisini betimleyeceğiz. Spesifik olarak, tekrarlı ölçümler deseninde katılımcılar tekrar tekrar test edildikleri için, değişirler. Örneğin, görevi daha iyi öğrendikleri için ya da deneysel ortamda rahatladıkları için, katılımcılar gelişme gösterebilirler. Ayrıca, denemelere bağlı olarak örneğin, yorgunluk ya da azalan motivasyon nedeniyle performansları da kötüleşebilir. Bu geçici değişikliklere *alıştırma etkileri* denir.

Bölüm 6'da açıklanan katılımcılar arasındaki bireysel farklar, seçkisiz gruplar deseninde ortadan kaldırılamaz; sadece seçkisiz atama yöntemi kullanılarak dengelenebilir. Aynı şekilde, tekrarlı ölçümler desenlerinde, tekrar tekrar test etmenin katılımcı üzerinde yarattığı alıştırma etkileri ortadan kaldırılamaz. Seçkisiz gruplar desenindeki bireysel farklar gibi, bir tekrarlı ölçümler desenindeki farklı koşullardan kaynaklanan alıştırma etkileri, yine de dengelenebilir veya ortalamaya yaklaştırılabilir. Alıştırma etkileri koşullar yoluyla dengelendiği zaman bağımsız değişken açısından karıştırıcı etki yaratmaz ve deney sonuçları yorumlanabilir.

Bu bölümde odaklandığımız esas nokta, araştırmacıların alıştırma etkilerini dengelemek için kullandıkları teknikleri açıklamaktır. Aynı zamanda, tekrarlı ölçümler desenleri için veri analizi işlemlerine başlayacağız ve tekrarlı ölçümler desenlerinde ortaya çıkabilen problemler üzerine bir değerlendirme ile bitireceğiz.

ARAŞTIRMACILAR NEDEN TEKRARLI ÖLÇÜMLER DESENLERİNİ KULLANIRLAR

- Araştırmacılar tekrarlı ölçümler desenlerini aşağıdaki nedenlerle kullanırlar: Denek sayısının az olduğu bir deneyi yürütmek, (2) bir deneyi daha etkin bir şekilde yürütmek, (3) deneyin duyarlılığını arttırmak ve (4) süreç içinde katılımcıların davranışlarındaki değişimleri incelemek.

Tekrarlı ölçümler desenini kullanmak, araştırmacılara birçok avantaj sağlar. İlk olarak, tekrarlı ölçümler desenleri bağımsız gruplar desenine göre daha az sayıda katılımcı gerektirir; bu nedenle az sayıda katılımcının olduğu deneylerde bu desenlerin kullanımı idealdir. Araştırmacıların çocuklarla, yaşlılarla ya da beyin hasarı olan katılımcılar gibi özel gruplarla yaptıkları çalışmalarda çoğunlukla, uygun katılımcı sayısı azdır.

Bağımsız gruplar desenlerinde yeterli sayıda katılımcı olduğunda bile araştırmacılar tekrarlı ölçümler desenlerini kullanmayı tercih ederler. Tekrarlı ölçümler desenleri sıklıkla daha uygun ve etkilidir. Örneğin, yüz tanımaya ilişkin ilginç bir çalışma düşünün. Yüzlere ilişkin belleğimizi etkileyen ön yargıları çok sayıda çalışma araştırmıştır. Araştırmalar kadınların erkek yüzlerine kıyasla kadın yüzlerini, bir başka yaş grubuna kıyasla kendi yaş gruplarındaki yüzleri, ve farklı ırktan yüzlere kıyasla aynı ırktan olan yüzleri daha iyi hatırladıklarını bulmuştur. Bu çalışmaların birçoğunda farklı yüz fotoğrafları katılımcılara sadece yarım saniye gibi çok kısa bir süre sunulur ve testte katılımcılara, fotoğraftakini daha önce görüp görmedikleri sorulur. Ayrı katılımcı gruplarına bir fotoğrafı bir saniyeden daha kısa bir sure göstermek çok yetersizdir. Bir deneysel koşula ait 20 fotoğrafın gösterildiği bir durumda bile, katılımcılara görevin niteliği hakkında yönerge vermek, görevin kendisini vermekten daha fazla zaman alır.

Temel Kavram

Tekrarlı ölçümler deseninin bir diğer önemli faydası, bağımsız gruplar deseninden genel olarak daha duyarlı olmasıdır. Bir deneyin **duyarlılığı,** etkisi küçük de olsa, bağımsız değişkenin etkisini belirleyebilme yeteneğidir. İdeal olarak, bir araştırmadaki katılımcıların bir deneysel değişimlemeye verdikleri tepkiler, birbirine benzer. Ama uygulamada, hepsinin aynı şekilde tepki vermediklerini biliyoruz. Bu *hata değişkenliği* (error variation), deney her yapıldığında uygulanan işlemlerdeki farklılıklara veya katılımcılar arasındaki bireysel farklılıklara bağlı olabilir. Bir deney koşulunda katılımcı tepkilerinde daha az değişkenlik varsa, bu deneyin daha duyarlı olduğu düşünülebilir ki bu da daha az hata değişkenliği demektir. Genel olarak, bir tekrarlı ölçümler desenindeki katılımcıların deney boyunca kendi içlerinde gösterdikleri değişim, bir seçkisiz gruplar desenindeki katılımcıların diğer katılımcılara göre gösterdikleri değişimden daha azdır. Bunu şöyle de ifade edebiliriz: Çoğunlukla insanların kendi *içinde* gösterdikleri değişim, birbirleri *arasında* gösterdikleri değişimden daha azdır. Bu nedenle hata değişkenliği, çoğunlukla tekrarlı ölçümler desenlerinde daha azdır. Daha az hata değişkenliği, bağımsız değişkenin etkisini daha kolay belirlemek demektir. Tekrarlı ölçümler desenlerinde bu şekilde artan duyarlılık, özellikle davranış üzerindeki küçük (gözlenmesi zor) etkileri olan bağımsız değişkenleri araştıran araştırmacılara çekici gelir.

Tekrarlı ölçümler desenleri bir başka nedenle daha kullanılır: Psikolojinin bazı alanlarında bu desene daha çok ihtiyaç duyulur. Araştırma sorusu, bir öğrenme deneyinde olduğu gibi, katılımcının davranışlarında zamanla ortaya çıkan değişimlerin anlaşılmasını içeriyorsa, bir tekrarlı ölçümler desenine ihtiyaç vardır. Dahası, her ne zaman deneysel işlem katılımcıların iki ya da daha fazla uyarıcıyı birbirlerine göre kıyaslamalarını gerektiriyorsa, bir tekrarlı ölçüm deseni kullanılmalıdır. Örneğin, romantik ilişkileri olan heteroseksüel çiftlerin, hızlı kararlarla insanların fotoğraflarını güvenilirlik, zekâ ve saldırganlık açısından derecelemeleri istendi (Günaydin, Zayas, Selcuk & Hazan, 2012). Katılımcılar farkında olmasa da yüzlerin yarısı kendi romantik arkadaşlarına benziyordu. Kadınlar kendi arkadaşlarına benzeyen yeni fotoğrafları, benzemeyen yeni fotoğraflara kıyasla daha olumlu değerlendirdiler. Erkeklerin kararları, romantik eşlerine olan benzerlikten etkilenmedi!

KUTU 7.1

TEKRARLI ÖLÇÜMLER VE TEKRARLI ÖLÇÜMLER DESENİ

Araştırmacıların katılımcıları tekrarlı bir şekilde test ettikleri farklı durumların birbirinden ayırt edilmesi önemlidir. Örneğin, Bölüm 5'te araştırmacıların, katılımcılarda zaman içinde ortaya çıkan değişiklikleri değerlendirmek için boylamsal araştırma deseninde aynı kişiye birden fazla uygulama yaptıklarını gördük. Tekrarlı ölçümler deseninde katılımcıların bir, iki ya da daha fazla koşuldaki davranışlarına ait ölçümleri karşılaştırmak için, bağımsız değişken değişimlenir. Önemli fark şudur: Tekrarlı ölçümler deseninde bir bağımsız değişken değişimlenir, ama boylamsal araştırma deseninde bu yapılmaz.

Tekrarlı test etme, araştırmacılar bir ölçümün güvenirliğini (tutarlılığını) araştırırken de kullanılabilir. Araştırmacılar testin test-tekrar test güvenirliği adı verilen güvenirliğini göstermek için, aynı bireylerden iki (ya da daha fazla) ölçüm alırlar (Bölüm 5'e bakınız). Sadece tekrarlı ölçümler deseni, katılımcı tepkilerinin farklı deneysel koşullarda karşılaştırıldığı bir bağımsız değişken içerir.

İnsanın işitsel ve görsel algısı ile duyusal süreçlerin fiziksel uyarıcılarla ilişkilerine dair araştırmalar (psikofizik), ağırlıklı olarak tekrarlı ölçümler desenlerini kullanır. *Perception & Psychophysics* ve *Journal of Experimental Psychology: Human Perception and Performance* gibi dergiler çoğunlukla tekrarlı ölçümler desenleri kullanan deney sonuçlarını yayımlarlar (Kutu 7.1'e de bakınız).

TEKRARLI ÖLÇÜMLER DESENLERİNDE ALIŞTIRMA ETKİLERİNİN ROLÜ

- Tekrarlı ölçümler deseninde bireysel farklar değişkeninin karıştırıcı etkisi gözlenmez, çünkü kişiler bağımsız değişkenin her koşulunda (düzeyinde) aynı katılımcıdır.
- Tekrarlı ölçümler deseninde katılımcıların performansı, bağımsız değişkenin etkisiyle değil, katılımcılar birden fazla test edildiği için değişebilir; buna, alıştırma etkileri denir.
- Bağımsız değişkenin farklı koşulları tüm katılımcılara aynı sırada uygulandığı zaman, alıştırma etkileri bir tekrarlı ölçüm deneyinin iç geçerliğini tehdit edebilir.
- Tekrarlı ölçümler deseninin iki şekli, tamamlanmış ve tamamlanmamış tekrarlı ölçümler deseni, alıştırma etkilerini farklı şekillerde kontrol eder.

Alıştırma Etkilerini Belirleme

Tekrarlı ölçümler desenlerinin şu ana kadar açıkladıklarımızın dışında, bir diğer önemli avantajı daha vardır. Bir tekrarlı ölçümler deseninde, katılımcıların bireysel özellikleri bağımsız değişkenin deneyde değişimlenmesi açısından karıştırıcı etki göstermez. *Aynı* katılımcılar tekrarlayıcı ölçümler desenlerinin tüm koşullarında test edilir; bu nedenle bir koşulun diğer koşula kıyasla daha

ŞEKİL 7.1 Yeni bir yeteneğin kazanılmasında alıştırma yapmanın hem olumlu hem de olumsuz etkileri vardır. Aynı deneyimin tekrarlanması gelişime yol açabilir ama aynı zamanda yorgunluk, motivasyonda azalma ve hatta sıkılmaya da yol açabilir.



parlak, daha sağlıklı ya da daha motive katılımcılardan oluşması olanaksızdır. Formel olarak belirtildiği gibi, *tekrarlı ölçümler desenlerinde bireysel farklar değişkeninin yarattığı karıştırıcı bir etki olmayabilir.* Bu yine de, tekrarlı ölçümler desenleri kullanılarak yapılan deneylerin iç geçerliğini tehdit eden hiçbir faktör olmadığı anlamına gelmez.

İç geçerliğe bir potansiyel tehdit, katılımcıların zaman içerisinde değişmeleridir. Tekrarlı ölçümler deseninde katılımcıların tekrarlı test edilmesi, onlara deneysel göreve ilişkin alıştırma yapma olanağı sağlar. Alıştırma yaptıkları için, katılımcılar verilen görevi yapmada gittikçe iyi ve daha iyi olurlar. Daha iyi olurlar, çünkü görev hakkında daha fazla şey öğrenirler ya da yorgunluk ve sıkılmaya bağlı olarak verilen göreve ilişkin giderek daha kötü performans gösterirler (Şekil 7.1'e bakınız). Tekrarlı ölçümler desenlerinde tekrarlı ölçüme bağlı olarak katılımcıların yaşadığı değişikliklere **alıştırma etkileri** (practice effects) denir. Genel olarak, tekrarlı ölçümler desenindeki alıştırma etkilerinin koşullar yoluyla dengelenmesi gerekir ki bu etki tüm koşullara "ortalama" dağılabilsin. Tekrarlı ölçümler desenlerinin kullanıldığı yorumlanabilir deneyler yürütmenin anahtarı, alıştırma etkilerini dengelemek için uygun tekniklerin kullanılmasını öğrenmektir.

Temel Kavram

Temel Kavram

Tekrarlı ölçümler deseninin iki şekli, tamamlanmış ve tamamlanmamış tekrarlı ölçümler desenleridir. Alıştırma etkilerini dengelemek için iki farklı teknik kullanılır ama bunların ikisi için de genel bir terim kullanılır: **Karşıt dengeleme** (counterbalancing). *Tamamlanmış desende*, alıştırma etkileri, koşullar her bir katılımcı için her seferinde farklı bir sırada uygulanarak dengelenir. Böylece her bir katılımcı, tamamlanmış bir deney olarak değerlendirilebilir. *Tamamlanmamış desende*, her koşul her katılımcıya *sadece bir defa* uygulanır. Koşulların uygulanma sırası, tamamlanmış desen durumunda olduğu gibi katılımcılar yoluyla değişir.

Tamamlanmış Desende Alıştırma Etkilerini Dengeleme

- Tamamlanmış desenlerde alıştırma etkileri, her katılımcının kendi içinde blok seçkisizleştirme ya da ABBA karşıt dengeleme tekniği kullanılarak dengelenir.
- Blok seçkisizleştirmede, deneyin tüm koşulları (bir blok), katılımcılara her uygulandığında, seçkisiz olarak sıralanır.
- ABBA karşıt dengeleme tekniğinde bütün koşullar seçkisiz şekilde verildikten sonra, bu sıranın tersi bir sırada verilir.
- Alıştırma etkisi doğrusal değilse veya katılımcıların performansı beklenti etkisinden etkilenebilecekse, ABBA karşıt dengeleme tekniği yerine seçkisiz bloklama tercih edilir.

Tamamlanmış desende her bir katılımcı için ortaya çıkabilecek alıştırma etkisini dengelemek için, katılımcılara yeterli sayıda uygulama yaptırılır. Görev yeterince basitse ve çok zaman harcamayı gerektirmiyorsa (örneğin, fotoğraflardaki insanların özellikleri hakkında yargıda bulunmak gibi), her bir katılımcıyı birkaç kez uygulamaya tâbi tutmak mümkündür. Aslında, bazı tamamlanmış desenlerde, bir ya da iki katılımcı test edilir ve katılımcıya yüzlerce uygulama yapılır. Yine de daha yaygın olan uygulama şekli bir kaç katılımcının test edilmesi ve her uygulamanın her katılımcıya nispeten az sayıda yapılmasıdır. Tamamlanmış bir desende araştırmacılar uygulamaların sırasına karar verirken iki seçenekleri vardır: Blok seçkisizleştirme ve ABBA karşıt dengeleme.

Blok seçkisizleştirme Bölüm 6'da seçkisiz gruplar deseninde, katılımcıları koşullara atamanın etkin bir yöntemi olarak blok seçkisizleştirme konusuna giriş yaptık. *Blok seçkisizleştirme*, tamamlanmış bir desende her bir katılımcı için koşulları sıralamak için de kullanılabilir. Katılımcılardan insanların fotoğraflarını güvenilirlik ya da sevimlilik gibi kişilik özellikleri açısından değerlendirmeyi gerektiren "İlk izlenimler" konusunda yapılan ve fotoğrafların çok çok kısa bir süre sunulduğu bir araştırmayı ele alın. Bu araştırmada fotoğraflar, milisaniye (msn) cinsinden üç farklı süre gösterildi: 100 msn, 500 ms ve 1,000 msn [Willis ve Todorov'a (2006) bakınız]. Bu kadar kısa sunum süreleri için gruplar arası deney deseni uygun olmazdı. Bu deneyde bütün katılımcılar, bu üç farklı sunum koşulunun her birisinde 22 yüz olmak üzere, toplam 66 yüzü değerlendirdiler. Bu deney her bir özellik için bir kez olmak üzere beş kez tekrarlandı ve farklı katılımcılar 66 resim için sadece bir kişilik özelliğini değerlendirdiler.

Blok seçkisizleştirme yöntemini kullanmak için, 66 deneme 22 blokluk üç denemeye ayrıldı. Her bir blokta, deneydeki üç koşulun her birisi, seçkisiz bir sırayla yer aldı. *Genel olarak, blok seçkisizleştirmede blokların sayısı her koşulun uygulanma sayısına ve bloğun büyüklüğü de deneydeki koşul sayısına eşittir.* Denemeler aşağıda gösterildiği gibidir:

Deneme	Koşul	
1	100	
2	1000	Birinci Blok
3	500	
4	1000	
5	100	
6	500	
.		
.		
.		
64	1000	
65	500	
66	100	

Blok seçkisizleştirme yönteminde katılımcıların dikkat, motivasyon veya fotoğrafları değerlendirme deneyimlerindeki farklılıklar, uyarıcının tüm sunum koşullarını aynı şekilde etkiler. Herhangi bir koşulun 66 denemenin başında, ortasında veya sonunda daha fazla görünmesi mümkün değildir. Alıştırma etkilerinin, üç deneysel koşulda (100, 500, 1000 msn maruz kalma süreleri) ortalama olarak aynı olmasını beklemek akla uygundur. Alıştırma etkilerinin kontrol edilmesinde blok seçkisizleştirmenin etkili olması için, her deneysel koşul bir kaç kez tekrar edilmelidir ki alıştırma etkilerinin ortalama şekilde dağılmalarını bekleyelim. Örneğimizde her koşul 22 kez ortaya çıkmaktadır. İki ya da üç blok sonra alıştırma etkilerinin dengelenmesini beklemeyiz –böyle bir beklenti, seçkisiz gruplar deseninde iki veya üç kişilik örneklemlerin denk olmalarını beklemekten farklı olmaz. Koşulları birden fazla sıralamak mümkün olmakla birlikte, tipik olarak tüm katılımcılar aynı blok seçkisizleştirme sıralamasını izler[1].

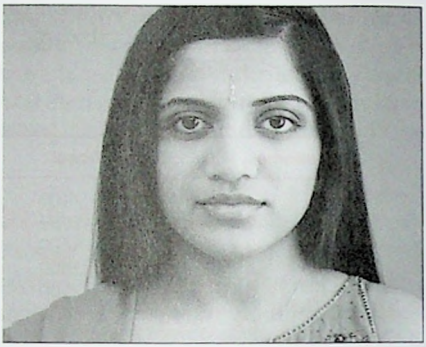
Willis ve Todorov'un (2006) deneyinde her denemeden sonra, ekranın ortasında 500 msn süreyle bir sabitleme noktası (+) belirdi. Katılımcılar, kısaca ekranda görünen fotoğraf hakkında yargıda bulundular (örneğin, "evet" ya da " hayır" diyerek bu insanın güvenilir olup olmadığını bildirdiler) ve bilgisayar

[1]Belki de bu şekilde betimlediğimiz Willis ve Todorov'un (2006) deneyinde, siz bir karıştırıcı etken fark ettiniz. 66 fotoğraf üç sete bölündü (her sette 22 fotoğraf olacak şekilde) ve her küme farklı bir görülme süresi koşuluna atandı (100, 500, 1000 msn). Bu işlem her bir katılımcının 22 fotoğraftan oluşan sete bakma süresini karıştırdı. Fotoğraf kümelerinden birinin 1000 msn'den daha yüksek bir hızda (örneğin, 50 msn) daha kolay değerlendirilmesi olasılığı var mıdır? Bunu bilmek zordur ama araştırmacılar işi şansa bırakmadı ve her fotoğraf seti katılımcıların üçte birine 50 msn, üçte birine 100 msn ve geriye kalan üçte birine de 1000 msn süreyle sunulacak şekilde dengelendi.

ŞEKİL 7.2 Willis ve Todorov'un (2006) sonuçlarına göre, bir insan hakkında ilk izlenim oluşturulabilmemiz için sadece 1/10 saniyeye ihtiyacımız vardır. Deneylerdeki fotoğraflarda gri tişört giymiş, sakal, dövme, küpe ya da gözlük gibi ayırt edici özellikleri olmayan aktörler yer almaktadır.



Bu insan güvenilir mi?



Bu insan yetenekli mi?

klavyesini kullanarak bu yargılarına ne kadar güvendiklerini belirttiler (7 aralıklı bir ölçek üzerinden). Altmış altı fotoğrafın hepsine tepkide bulunmak 15 dakikadan daha az zaman aldı.

Biz bu araştırmacıların blok seçkisizleştirme yöntemini kullandığını söylemiş olsak da aslında araştırmacılar 66 fotoğrafı her katılımcıya bilgisayarın ürettiği yeni bir sırada seçkisiz olarak sundular. Basit bir seçkisizleştirme işlemi, blok seçkisizleştirme yöntemindeki kadar kontrollü değildir ama biz, özellikle sunum zamanları kısa olduğunda, alıştırma etkilerinin genel olarak seçkisiz dizilimler arasında ortalama şekilde dağılacağını varsayabiliriz. Çok sayıda seçkisiz dizi kullanıldığında, bu varsayım güçlenir. Şimdi neden böyle olduğunu açıklayalım.

Blok seçkisizleştirme, deneyin her bir koşulunun bütün aşamalarda (yani başında, ortasında, sonunda) eşit sayıda sunulmasını garanti eder. Bununla birlikte, basit bir seçkisizleştirme işlemi kullanıldığında, uygulamanın bir aşamasındaki bir koşulun, bir diğer koşula kıyasla, *tesadüfen* daha fazla ortaya çıkması olasılığı vardır. Eğer böyle olursa, bir katılımcı için koşullar aracılığıyla alıştırma etkilerinin dengelendiğini varsayamayız. Birkaç seçkisiz dizi kullanıldığında, bir koşulun başlangıçta, ortada ve denemelerin sonunda daha sık ortaya çıkma olasılığı daha azdır. Dahası, sonuçlar deneydeki tüm katılımcılar vasıtasıyla özetleniyorsa, alıştırma etkilerindeki herhangi bir farklılık dengelenmelidir. Bu varsayım, daha sonra üzerinde duracağımız, tamamlanmamış tekrarlı ölçümler deseninde yapılana benzerdir. Bu yüzden, gerçek bir *blok* seçkisizleştirme işleminden farklı olarak, bir deneydeki koşulların sunumunda dizilerin basit seçkisizleştirmesi kullanılacaksa, çok sayıda farklı seçkisiz sıra kullanılmalıdır.

Esas nokta şudur: Gerçek bir blok seçkisizleştirme işlemindeki değişkenlikler, alıştırma etkisi deneydeki koşullar yoluyla dengelenmişse, kullanılabilir. Öte yandan, bir blok seçkisizleştirme işlemi bu iddiayı güçlendirir.

TABLO 7.1 ÜÇ SUNUM KOŞULU OLAN (HIZLI, ORTA, YAVAŞ) BİR DENEYDE ABBA'YA GÖRE DENGELENMİŞ DENEMELER DİZİSİ

	Deneme 1	Deneme 2	Deneme 3	Deneme 4	Deneme 5	Deneme 6
Koşul:	Orta	Hızlı	Yavaş	Yavaş	Hızlı	Orta
Alıştırma etkisi (doğrusal)	+0	+1	+2	+3	+4	+5
Alıştırma etkisi (doğrusal olmayan)	0	+6	+6	+6	+6	+6

Yaptıkları deneyin sonuçları şunu gösterdi: Sunum süresinde 1/10 saniyelik (100 msn) bir artış, beş kişilik niteliği (çekicilik, sevimlilik, güvenilirlik, yeteneklilik ve saldırganlık) hakkında verilen hükümler için katılımcılar ile zaman kısıtlaması yapılmayan katılımcıların değerlendirmeleri arasındaki korelasyonu arttırmadı. Bir insan hakkında izlenim oluşturmak için saniyenin onda biri kadar zaman yeterlidir. (Şekil 7.2'ye bakınız). Unutulmamalıdır ki bu, ilk izlenimin doğru olduğu anlamına gelmez [Bölüm 2'de tartışılan Rule ve arkadaşlarının (2013) bulgularına bakınız].

ABBA Karşıt Dengeleme En basit şeklinde ABBA karşıt dengeleme, tamamlanmış desende her bir koşul en az iki defa uygulanarak alıştırma etkilerini dengelemek için kullanılabilir. ABBA karşıt dengelemede koşullar seçkisiz sunulduktan sonra bu sıranın tersi uygulanır. Bu nedenle, tekrarların çift olması gerekir. Deneyde iki koşul (A ve B) varsa, bu teknikte ne yapılacağı bellidir ama ABBA karşıt dengeleme, sadece iki koşullu deneylerle sınırlı değildir. Üç koşulun olduğu bir deneyde, denemelerin dizilişi ABCCBA şeklinde olabilir. Üç koşulun, ilk üç denemedeki sıralaması, 4, 5 ve 6'ncı denemelerin tersidir. (Tablo 7.1'e bakınız).

ABBA karşıt dengeleme işleminin bazı sınırlıkları, onu birçok durumda ideal olmaktan uzaklaştırır. Örneğin, ABBA karşıt dengeleme sadece alıştırma etkileri doğrusal olduğu zaman uygun şekilde kullanılabilir. Eğer alıştırma etkileri doğrusal ise aynı miktarda alıştırma etkileri ardışık her denemede performansa eklenir veya çıkarılır. Tablo 7.1'deki "Alıştırma etkileri (doğrusal)" satırı, ABBA karşıt dengeleme tekniğinin alıştırma etkilerini nasıl dengelediğini gösteriyor. Bu örnekte, varsayımsal (hipotetik) alıştırma etkilerinin bir *birimi*, her denemedeki performansa eklenir. İlk denemeyle ilgili hiç bir alıştırma etkisi olmadığı için, tablodaki Deneme 1'e eklenen alıştırma miktarı sıfırdır. Deneme 2'de, katılımcıların ilk denemedeki deneyimleri nedeniyle, 1 birimlik varsayımsal etki vardır; Deneme 3'te, katılımcıların iki denemedeki deneyimleri nedeniyle eklenmiş iki birimlik alıştırma etkisi vardır ve bu böyle devam ediyor.

Alıştırma etkileri doğrusal olduğunda, alıştırmanın varsayımsal "birimleri" tüm koşullar için aynıdır (Tablo 7.1'e bakınız). Üç koşulda, her koşul için 5 birim kazanılmıştır. Örneğin, "Orta" koşulu, toplam 5 birimlik alıştırma etkisinde en küçük (0) ve en büyük (5) alıştırma etkisini almaktadır; bu örneğimizde varsayımsal birimlerin toplamı Hızlı (4 + 1) ve Yavaş (3 + 2) koşulları için de 5'tir.

Şimdi Tablo 7.1'de bir alttaki *Alıştırma etkisi (doğrusal olmayan)* satırını ele alın. Birinci denemeden ikinci denemeye değişim, 0'dan 6'ya çıkmıştır; bu, katılımcıların sadece bir denemede büyük bir değişime uğradığını gösteriyor. Ne yazık ki bunun ne zaman olabileceğini görmek her zaman kolay değildir. Bu örnekte, *Yavaş, Orta ve Hızlı* koşullarla ilişkili alıştırma miktarı dengelenmemiştir (sırasıyla, +6, +12 ve +12). Asıl verileri toplamaya başlamadan önceki çeşitli "ısınma" denemeleri bazen alıştırma etkilerindeki ani değişiklikleri sona erdirebilir.

ABBA karşıt dengelemenin bir diğer sınırlılığı, beklenti etkileri ortaya çıkınca görülmektedir. Aşağıdaki koşullar diziliminin size verildiğini varsayın:

Orta Hızlı Yavaş Yavaş Hızlı Orta Orta Hızlı Yavaş Yavaş Hızlı Orta
Orta Hızlı Yavaş Yavaş Hızlı Orta

Bir katılımcı bir sonraki sırada hangi koşulun ortaya çıkacağına dair beklenti oluşturduğu zaman *beklenti etkileri* ortaya çıkar; siz yukarıdaki dizilimle karşılaştığınız zaman böyle bir beklenti oluşturdunuz. Bir katılımcının bir koşula tepkisi, koşulun kendisine ait gerçek deneyimden çok, beklentilerden etkilenebilir. Örneğin Yavaş koşulundan sonra "gevşeyen" bir katılımcının, hemen ikinci bir Yavaş koşulu beklentisine girdiğini düşünebiliriz.

Beklenti etkilerinin oluşmasını önlemek için araştırmacılar, genellikle sadece koşul sayısı ve her koşulun tekrar sayısı göreceli küçük olduğunda ABBA karşıt dengeleme tekniğini göz önüne alırlar. Ama bu yine de doğrusal özellik göstermeyen alıştırma etkilerine ilişkin potansiyel problemleri ortadan kaldırmaz. Bu endişeler ve bir ABBA işleminde koşulların iki defa tekrarlanmasının gerekli olması, tamamlanmış tekrarlı ölçümler desenlerinin çoğunda, blok seçkisizleştirmeyi daha iyi bir karşıt dengeleme alternatifi hâline getirir. Bir blok seçkisizleştirme işleminde, eğer deney sadece bir kaç koşulun tekrarını gerektiriyorsa, araştırmacı birden fazla seçkisiz koşullar dizilimi kullanmalıdır.

Tamamlanmamış Desende Alıştırma Etkilerini Dengeleme

- Tamamlanmış desende alıştırma etkileri, her bir denek için dengelenir oysa tamamlanmamış desende bu etkiler, denekler *arasında* yapılır.
- Tamamlanmamış desende alıştırma etkilerini dengelemenin kuralı şudur: Deneyin her bir koşulu, her bir sırada (birinci, ikinci vs.) eşit sayıda sunulmalıdır.
- Dört veya daha az koşulu olan tamamlanmamış desende, alıştırma etkilerini dengelemenin en iyi yöntemi, koşullara ait mümkün tüm sıralamaları kullanmaktır.
- Tamamlanmamış bir desende spesifik dizilimleri seçmede kullanılan iki yöntem Latin karesi ve seçkisiz başlayan dönüşümlü dizilim'dir (random starting order with rotation).
- İster olası tüm sıralar isterse seçilmiş sıralar kullanılsın, katılımcılar farklı sıralamalara seçkisiz olarak atanmalıdır.

Tamamlanmamış desende, her uygulama her katılımcıya *sadece bir defa* verilir. Bu yüzden herhangi bir katılımcıya ait sonuçlar yorumlanamayabilir, çünkü bağımsız değişkenin her bir katılımcıya uygulanan düzeyi, eşit olasılıkla bütün düzeylerin herhangi birisi olabilir. Örneğin, bir tamamlanmamış desen

deneyindeki ilk katılımcı önce deneysel koşulda (E) ve ikinci sırada da kontrol koşulunda (C) test edilebilir. Deneysel ve kontrol koşulları arasında gözlenen performans farkı, bağımsız değişkenin etkisine *veya* EC sıralamasından kaynaklanan alıştırma etkilerine bağlı olabilir. Koşulların sırasının ve bağımsız değişkenin yarattığı bu karıştırıcı etkiyi önlemek için, farklı katılımcılara farklı koşul sıralamaları uygulayabiliriz. Örneğin, yukarıdaki tamamlanmamış desen deneyimizde ikinci katılımcıya önce kontrol koşulunu, sonra deneysel koşulu vererek CE sıralaması uygulayabilirdik. Bu şekilde, bir yerine iki katılımcı kullanarak iki koşul arasındaki sıralama etkilerini dengeleyebilirdik.

Tamamlanmamış tekrarlı ölçümler deseninde, etkileri dengelemede kullanılan teknikleri örneklerle açıklamak için, damgalama (stigma) alanına ait literatürü inceleyebiliriz. Bu araştırmalar, toplum tarafından damgalamanın madde, alkol ve kumar gibi bağımlılıklar ile akıl hastalıkları gibi konularda insanların yardım alma istekleri üzerindeki etkilerini incelemektedir. "Zihinsel sağlığının" bozuk olması korkusu, psikolojik sorunlar yaşayan bir çok insanı yardım istemekten alıkoymaktadır (Şekil 7.3). Araştırmacılar çoğunlukla genel ve kişisel olmak üzere iki tip damgalamadan söz etmektedir (Corrigan, 2004). Toplumsal damgalar olumsuz kalıp yargıları ifade eder ve toplumda akıl hastalığı olan bireylere yönelik ön yargıların gelişmesine yol açar. (örneğin," Akıl hastalığı olan insanlar tehlikelidir" inancı gibi). Kişisel damga, bir bireyin kendi kişisel tutumlarını işaret eder; bireyin toplumsal damgalamayı içselleştirdiği zaman yaptığı gibi (örneğin, "Zihinsel rahatsızlığım için tedavi arasaydım kendimi aşağılanmış hissederdim." gibi). Damgalama üzerine araştırmaların çoğu, anket yöntemi kullanılarak yapılmıştır Örneğin,

ŞEKİL 7.3 Kumar oynama, madde ve alkol bağımlılığı gibi bir çok psikolojik rahatsızlığa yönelik toplumsal damgalama, bireyleri yardım istemekten alıkoyan önemli bir engeldir.



5000'den fazla kolej öğrencisi ile internette yapılan ankette, kişisel damgalamanın kadınlara kıyasla erkeklerde ve beyaz öğrencilere kıyasla siyah ve Asyalı öğrencilerde daha yüksek olduğu bulunmuştur (Eisenberg, Downs, Golberstein & Zivin, 2009).

Akıl hastalığına yönelik damgalamayı araştıran bir diğer temel yaklaşım, psikolojik rahatsızlığı olan kişileri betimleyen kısa hikâyelerin okunması yöntemini kullanmaktadır. Katılımcılar daha sonra öfke, korku, güvensizlik ve kaçınma gibi duygularını değerlendirirler ve sorulara muhtemelen, o kısa hikâyede bahsedilen kişi ile kendileri arasında kurdukları yakınlıkla uyumlu cevaplar verirler. Bazen bir araştırmada sadece bir hikâye kullanılır (örneğin, Bathje & Pryor, 2011). Bununla birlikte, çoğunlukla katılımcıların bir kaç kısa hikâyeye cevap vermesi, araştırmacıların farklı hastalıklara karşı tutumları kıyaslamalarına olanak sağlar. Bir çalışmada katılımcılar anoreksiya blumia (aşırı yeme rahatsızlığı), şişmanlık veya ağır depresyonu olan kadınları betimleyen beş kısa hikâye okudular. Katılımcılar, yeme bozukluğu olanları, depresyonlu kadınlardan daha fazla suçladılar ve bu durumdan en çok obez kadınların kendilerinin sorumlu olduklarını savundular (Ebneter & Latner, 2013).

Katılımcılardan bir kısa hikâyeden fazlasına cevap vermelerini istemek, veri elde etmek için etkili bir yoldur ve farklı özellikleri olan hedefleri kıyaslama olanağı sağlar. Örneğin, Kanadalı lise öğrencilerinin kumar oynama sorununa ilişkin damgalamalarını inceleyen araştırmacılar, kıyaslama yapmak için araştırmalarında kanserli, şizofren ve alkol bağımlısı bireylerin kısa hikâyelerini kullandılar (Horch & Hodgins, 2008). Katılımcılar tüm hikâyeleri sadece bir kez gördükleri için, kısa hikâyelerin kullanıldığı çalışmalar bir tamamlanmamış tekrarlı ölçümler desenine örnektir.

Tamamlanmamış bir desende koşulların sunum sırasının değiştirilerek alıştırma etkilerinin dengelenmesi zorunludur. Bu desende alıştırma etkileri, basit bir kuralla dengelenir: *Deneyin her bir koşulu her bir sırada (1'inci 2'nci, 3'üncü, 4'üncü sırada vb. gibi) eşit sıklıkta yer almalıdır.* Bu genel kuralı uygulamanın uygun birkaç tekniği vardır. Bu teknikler sağladıkları dengeleme şekline göre ayrılırlar; temel kuralın karşılanması ve deneyin yorumlanabilmesi, ancak bu tekniklerin uygun şekilde kullanılmasıyla mümkündür. Bu demektir ki eğer uygun dengeleme tekniği kullanılırsa, davranışı etkileyen faktörün alıştırma etkisi mi yoksa bağımsız değişken mi olduğunu belirleyebiliriz.

Olası Bütün Dizilimler Tamamlanmamış desende alıştırma etkilerini dengelemek için tercih edilen teknik, tüm olası koşullara ilişkin dizilimleri kullanmaktır. Her katılımcı seçkisiz olarak dizilerden birine atanır. Sadece iki koşulun olduğu durumda, sadece iki olası dizilim vardır (AB ve BA); üç koşul için 6 olası dizilim vardır (ABC, ACB, BAC, BCA, CAB, CBA). Genel olarak, N sayıdaki koşul için, mümkün N! (ki "N faktöryel" olarak okunur) sıra vardır. Burada N!, $N(N - 1)\ (N - 2)\ \ldots\ (N - [N - 1])$'e eşittir. Görüldüğü gibi üç koşul varsa, altı mümkün dizilim vardır: $3!\ (3 \times 2 \times 1 = 6)$. Koşulların sayısı arttıkça, ihtiyaç duyulacak dizilimlerin sayısı da çok hızlı bir şekilde artar. Örneğin, beş koşul için 120 olası dizilim ve altı koşul için 720 olası dizilim vardır. Bu nedenle, olası dizilimlerin kullanımı, dört ya daha az koşulu içeren deneylerle sınırlandırılmıştır.

TABLO 7.2 DÖRT KOŞULLU TAMAMLANMAMIŞ BİR TEKRARLI ÖLÇÜMLER DESENİNDE, ALIŞTIRMA ETKİLERİNİ DENGELEMEK İÇİN ALTERNATİF TEKNİKLER

Olası Tüm Dizilimler								Seçilmiş Dizilimler							
								Latin Karesi				Seçkisiz Başlayan Dönüşümlü Sıra			
Dizilimdeki Yeri				Dizilimdeki Yeri				Dizilimdeki Yeri				Dizilimdeki Yeri			
1'inci	2'nci	3'üncü	4'üncü	1'inci	2'nci	3'üncü	4'üncü	1'inci	2'nci	3'üncü	4'üncü	1'inci	2'nci	3'üncü	4'üncü
KO	Ş	K	N	K	KO	Ş	N	KO	Ş	K	N	Ş	K	N	KO
KO	Ş	N	K	K	KO	N	Ş	Ş	N	KO	K	K	N	KO	Ş
KO	K	Ş	N	K	Ş	KO	N	N	K	Ş	KO	N	KO	Ş	K
KO	K	N	Ş	K	Ş	N	KO	K	KO	N	Ş	KO	S	K	N
KO	N	Ş	K	K	N	KO	Ş								
KO	N	K	Ş	K	N	Ş	KO								
Ş	KO	K	N	N	KO	Ş	K								
Ş	KO	N	K	N	KO	K	Ş								
Ş	K	KO	N	N	Ş	KO	K								
Ş	K	N	KO	N	Ş	K	KO								
Ş	N	KO	K	N	K	KO	Ş								
Ş	N	K	KO	N	K	Ş	KO								

Not: Damga deneyindeki dört koşul kısa hikayenin türünü göstermektedir: Kumar Oynama (KO), Şizofreni (Ş), Kanser (K) ve Nötr (N).

Damgalama çalışan araştırmacıların alıştırma etkilerini nasıl dengeleyebileceklerini ele alın. Varsayalım ki kumar oynama sorunu, kanser, şizofreni veya klinik olmayan hafif bir stresi (nötr koşul) olan bireyleri betimleyen dört kısa hikâye var. Tüm olası koşulların kapsanması için toplam 24 dizilim gerekir. Bu dizilimler (koşulların sırası) Tablo 7.2'nin sol yarısında sunulmuştur. Tüm olası dizilimleri kullanmak, tüm koşulların her sırada eşit sayıda görünmesini sağlayan genel kuralı karşılar. İlk dizilim, bu dengelemeyi en açık şekilde gösteriyor: İlk altı dizilim, kumar oynama (KO) koşulu ve sonraki altı dizilimin her birisi şizofreni (Ş), kanser (K) ve nötr (N) koşullarla başlıyor. Bu örüntü aynen, sıralı dört konumun her birisine uygulanır. Örneğin, KO koşulu aynı zamanda ikinci, üçüncü ve dördüncü sıralı konumun her birinde altı kez ortaya çıkar. Aynısı Ş, K ve N koşulları için de doğrudur.

Olası tüm dizilimleri kullanmaya karar verirken üzerinde durulması gereken bir başka konu daha vardır. Bu tekniğin etkin olması için en az bir katılımcının, koşulların mümkün dizilimlerinin her birisinde test edilmesi gerekir. Bu da en az bir katılımcının KO-Ş-K-N, yine en az bir katılımcının KO-Ş-N-K ve bunun gibi devam eden dizilimleri alması demektir. Bu yüzden tüm dizilimlerin uygulanabilmesi için, en az olası tüm dizilimlerin sayısı kadar da katılımcı olması gerekir. Yani, deneyde eğer dört koşul varsa, en az 24 katılımcı gerekir (ya da 48, ya da 72, ya da 24'ün diğer çarpanları kadar). Bu sınırlama nedeniyledir ki ilk katılımcıyı test etmeden önce, araştırmacının çalışmaya katılabilecek katılımcı sayısını bilmesi gerekir.

Seçilmiş Diziler Tamamlanmamış desende tüm olası koşullarda alıştırma etkilerini dengelemek için tercih edilen yöntemi betimledik. Bununla birlikte olası tüm dizilerin kullanımının pratik olmadığı durumlar vardır. Örneğin, tamamlanmamış bir desende yedi düzeyi olan bir bağımsız değişkeni kullanmak isteseydik, ve bunun için tüm olası koşulları kullanmış olsaydık, yedi koşula (7! dizi) ait olası dizilerin her birisi için bir katılımcı ayıracaktık; böylece, 5040 katılımcıyı test etmemiz gerekecekti. Beş veya daha fazla koşulu olan deneylerde tamamlanmamış desen kullanmamız gerekiyorsa, açıktır ki tüm olası dizilimleri kullanmak için bazı alternatiflere ihtiyacımız vardır.

Alıştırma etkileri, tüm olası koşulların sadece bazılarını kullanarak dengelenebilir. Seçilmiş dizilerin sayısı daima deneydeki koşul sayısının bir kaç katına eşit olacaktır. Örneğin, yedi düzeyli bir bağımsız değişkenle bir deneyi yaparken, alıştırma etkilerini dengelemek için, yedi dizilimin 7, 14, 21, 28 ya da diğer katlarını seçme ihtiyacı duyarız. Seçilmiş dizileri kullanmanın iki temel şekli Tablo 7.2'dedir. Dengeleme türlerini daha doğrudan bir biçimde karşılaştırabilmeniz için, önceki bölümde açıkladığımız damgalama deneyindeki dört düzeyli bağımsız değişkene ait seçilmiş diziler tekniklerini örneklerle açıkladık.

Seçilmiş dizileri kullanan dengelemenin ilk türü Latin Karesidir. *Latin Karesi*, alıştırma etkilerinin genel kuralını yerine getirir. Bu kural, her koşulun her sıralı konumda bir kez ortaya çıkmasıdır. Örneğin, Tablo 7.2'nin merkezinden sağa doğru baktığımızda, Latin Karesi koşulunda "KO"nın birinci, ikinci, üçüncü ve dördüncü sıralarda bir kez ortaya çıktığını görürüz. Bu her koşul için böyledir. Ayrıca, Latin Karesinde her koşul en az bir defa diğerinden önce ve sonra gelir. Nitekim, Tablo 7.2'deki Latin Karesinde "Ş-KO" ve "KO-Ş" dizilimleri en az birer defa ortaya çıkmıştır. "K-Ş" sıralaması ve "Ş-K" sıralaması da en az birer kez ortaya çıkmıştır ve bu kural, her koşul sıralaması için böyle devam etmektedir (Latin karesi oluşturma tekniği Kutu 7.2'de açıklanmıştır).

Seçilmiş dizileri kullanan ikinci dengeleme tekniği, bir seçkisiz koşullar dizisi ile başlanmasını ve her seferinde her koşulu bulunduğu pozisyondan sola kaydıracak şekilde dizilimi sistematik olarak döndürmeyi gerektirir (Tablo 7.2.'nin sağ tarafındaki örneğe bakınız). Döndürme yolu ile başlangıçta seçkisiz bir dizinin kullanımı, etkin bir şekilde alıştırma etkilerini dengeler, çünkü Latin Karesinde olduğu gibi, her koşul her sırada ortaya çıkar. Bununla birlikte, dizilimlerin sistematik şekilde döndürülmesi, her koşulun daima *aynı* diğer koşullardan önce ve daima sonra geldiği anlamına gelir (örneğin, N daima C'den sonra ve G'den önce gelir) ama bu, Latin karesi tekniğinden farklıdır. Bu yöntemin temel avantajları, döndürme tekniği ile seçkisiz bir başlangıç yapmanın kolaylığı, ve koşul sayısı dörtten fazla olan deneylerde uygulanabilmesidir.

Tüm olası dizilimlerin, Latin Karesinin ve seçkisiz başlayan dönüşümlü dizilimlerin kullanımı, alıştırma etkilerinin dengelenmesi için aynı düzeyde etkilidir, çünkü bu üç tekniğin üçü de her bir koşulun her sırada eşit sıklıkta yer almasını sağlar. Alıştırma etkilerini dengelemek için hangi tekniğin kullanıldığından bağımsız olarak, koşulların dizilimi tam anlamıyla ilk katılımcının test edilmesinden önce belirlenir ve katılımcılar seçkisiz olarak bu dizilimlere atanır.

KUTU 7.2

LATİN KARESİ NASIL OLUŞTURULUR

1 Deneydeki koşulları seçkisiz olarak sıralayın.

2 Seçkisiz sıraladığınız bu koşulları, 1'den başlayarak N'e kadar numaralayın. Böylece, eğer 4 koşulunuz varsa (N = 4; A, B, C, D) ve baştaki seçkisiz diziliminiz B, A, D, C şeklinde olduysa, bu durumda B = 1, A = 2, D = 3, C = 4 olur.

3 İlk sırayı (koşulların ilk dizilimini) oluşturmak için şu kuralı uygulayın: 1, 2, N, 3, N-1, 4, N-2, 5, N-3, 6 ve bu şekilde devam edin. Bizim örneğimizde bu dizilim 1, 2, 4, 3 şeklinde olur.

4 İkinci sırayı (koşulların ikinci dizilimini) oluşturmak için, ilk dizilimdeki her bir sayıya 1 ekleyin ama N=1'e 1 eklediğinizin farkında olarak ekleyin. Böylece dizilimimiz 2, 3, 1, 4 olur.

5 Üçüncü sıra (koşulların üçüncü dizilimi) ikinci dizilimdeki her sayıya 1 ekleyerek ve tekrar N+1=1 olacak şekilde ilerleyin; böylece üçüncü dizilim 3, 4, 2, 1 olur.

6 Sonraki her sıra için bu işlem uygulanır. Dördüncü sırayı da siz oluşturabilir misiniz?

7 İkinci basamakta yapıldığı gibi, siz de aşağıdaki koşullara numaralar verin. Bu örneğin Latin Karesi şöyle olur.

B A C D

A D B C

D C A B

C B D A

Eğer koşullar tek sayılıysa, o zaman iki kare oluşturulmalıdır. İlk dizilim, yukarıda çift sayılı karelere uygulanan kurala göre yapılabilir. İkincisi, birinci karedeki diziler tersine çevrilerek oluşturulabilir. Örneğin, varsayın ki N = 5 ve ilk karenin ilk dizilimi B A E C D'dir. Bu durumda ikinci karenin ilk dizilimi D C E A B olur. Bir N X 2N karesi oluşturmak için iki kare birleştirilir. Tek veya çift sayılı olsun, her iki durumda da denekler karenin sıralarına seçkisiz atanmalıdır. Bu nedenle, en az sıraların katları kadar katılımcınız olmalıdır [Latin Kareleri oluşturma teknikleri aynı zamanda Winer, Brown, ve Michels'de de (1991,s. 674–679) betimlenmiştir.].

TEKRARLI ÖLÇÜMLER DESENLERİNDE VERİ ANALİZİ

Sonuçların Betimlenmesi

- Tamamlanmış desenin veri analizine, her katılımcı için bir özet puan (örneğin; ortalama, medyan) hesaplaması yapılarak başlanır.
- Betimleyici istatistik, bütün katılımcıların bağımsız değişkenin her bir koşulundaki performansını özetlemek için kullanılır.

Veriyi hatalar ve uç değerler açısından kontrol ettikten sonra, tekrarlı ölçümler desenindeki bir analizin ilk aşaması, katılımcıların deneyin her koşulundaki performansını özetlemektir. Seçkisiz gruplar deseninde bu, deney koşullarının her birinde test edilen katılımcıların puanlarını listelemek ve sonra ortalama ve standart sapma gibi betimleyici istatistiklerle bu puanları özetlemek demektir. Bir tamamlanmamış tekrarlı ölçümler deseninde her katılımcı her koşulda bir puan alır ama bu desen her bir koşuldaki puanları özetlemek için hâlâ basittir.

Tüm puanlar her koşul için bir kez birlikte listelendiğinde, ortalamalar ve standart sapmalar her koşulda performansı betimlemek için hesaplanabilir.

Bir tamamlanmış tekrarlı ölçümler desenini analiz ederken, bir adım daha atılması gerekir. Sonuçları özetlemeden ve betimlemeye başlamadan önce, her koşuldaki her katılımcı için bir puan hesaplamanız gerekir. Bu adım gereklidir, çünkü tamamlanmış desende her katılımcı her koşulda birden fazla test edilir. Örneğin, tamamlanmış tekrarlı ölçümler desenine göre tasarlanmış bir zaman algısı deneyinde, sınıfa benzetilmiş bir ortamda beş katılımcının test edildiklerini varsayın. Deneyin amacı katılımcıların zaman tahminlerinin, gerçek zaman aralıklarına göre doğruluğunu test etmek değildi. Amaç, zaman aralıkları arttıkça katılımcıların zaman tahminlerinin de sistemli şekilde artıp artmadığını belirlemekti. Bir başka ifadeyle, acaba katılımcılar farklı zaman aralıkları arasında ayrım yapabiliyor muydu?

Deneydeki her katılımcı dört farklı zaman aralığının (12, 24, 36 ve 48 sn) her birinde altı kez test edildi. Zaman aralıklarında sunulan dizileri belirlemek için blok seçkisizleştirme tekniği kullanıldı. Böylece her katılımcı dört zaman aralığı uzunluğunun her birisi için altı tahmin olmak üzere toplam 24 zaman tahmini yaptı. Verilen bir zaman aralığı için yapılan altı değerlendirmenin her biri alıştırma etkileriyle bozulmaktadır; bu nedenle, altı değerlendirmeye ait bilgileri birleştirirken bazı ölçülere ihtiyaç vardır. Tipik olarak, her koşulda tek bir performans tahmini elde etmek için, her bir zaman aralığı için altı tahminden elde edilen ortalama hesaplanır. Ama hatırlayacağınız gibi, yine de uç değerler ortalamayı etkileyebilir; her bir zaman aralığındaki altı testten en az birisi için, katılımcıların uç tahminler yapması olasıdır. Bu nedenle, bu özel veri kümesi için, belki de katılımcıların zaman aralıklarına ilişkin tahminlerini ifade eden en iyi ölçüm, altı tahminin ortanca (medyan) değeridir. En yakın tam sayılara yuvarlanan bu medyan tahminleri Tablo 7.3'tedir (*Grup* performansını özetleyen betimleyici istatistikler olarak, ortalama ve medyanı görmeye alışık olabilirsiniz; yine de, örnekte açıklandığı gibi, "deneme" ya da "testler" üzerinden elde edilen bir "ortalama" olduğunda, bu özetleyici istatistik aynı zamanda bir *kişinin* performansını göstermek için de kullanılır.).

TABLO 7.3 BİR TEKRARLI ÖLÇÜMLER DESENİ DENEYİNE AİT VERİ MATRİKSİ TABLOSU

Veri Matriksi				
	Zaman Aralığının Uzunluğu			
Katılımcı	12	24	36	48
1	13	21	30	38
2	10	15	38	35
3	12	23	31	32
4	12	15	22	32
5	16	36	69	60
Ortalama (S)	12.6 (2.0)	22.0 (7.7)	38.0 (16.3)	39.4 (10.5)

Not: Tablodaki zaman aralığı uzunluğu değişkeninin her düzeyinde katılımcıların altı tepkisinin medyanı görülmektedir. Alt sıradaki ortalamalar, medyanların (her zaman aralığı uzunluğunda beş katılımcıdan alınan altı tepkinin) ortalamalarıdır.

Bir defa her katılımcı için her bir koşulda bir puan elde edildikten sonraki adım, tüm katılımcılara ait sonuçları uygun betimsel istatistiklerle özetlemektir. Dört zaman aralığının her birisi için tahminî ortalama ve standart sapma, Tablo 7.3'te "Ortalama (S)" olarak belirtilen sıradadır. Tabloda sadece beş katılımcı yer alıyor olsa da bu tahminî ortalamalar katılımcıların farklı uzunluktaki zaman aralıkları arasında ayrım yapmış görünmektedir; bu, en azından 36 saniyeye kadar olan zaman aralıkları için böyledir.



Bölüm 6'da bahsettiğimiz gibi, bir deneyin sonuçlarını betimlerken, etki büyüklüğü ölçümlerini de bildirmek uygun olur. Tekrarlı ölçümler deseni ile ilişkili tipik bir etki büyüklüğü ölçüsü, eta kare (eta squared – η^2) adı verilen ölçüdür. Zaman algısı deneyine ait eta kare değeri .80 idi. Bu değer katılımcıların zaman tahminlerindeki değişkenliğin büyük bir kısmının, bağımsız değişkenle (zaman aralığının uzunluğuyla) açıklanabileceğini göstermektedir. Bölüm 11 ve 12'de etki büyüklüklerinin hesaplanmasına ve yorumlanmasına ilişkin daha fazla bilgi bulabilirsiniz. Bölüm 12'de, Tablo 7.3'teki veriler kullanılarak eta kare değerinin nasıl hesaplanacağı açıklanacaktır.

Verilerin Anlamını Onaylama

- Tekrarlı ölçümlere ilişkin genel işlemler ve sıfır hipotezini test etme ve güven aralıklarının mantığı, seçkisiz gruplar desenleri için kullanılanlara benzerdir.

Tekrarlı ölçümler desenlerini kullanan deneylerdeki veri analizi, Bölüm 6'da bahsedilen seçkisiz gruplar deseni deneylerindeki işlemlerin aynısını içerir. Araştırmacılar bağımsız değişkenin davranış üzerinde bir etkisi olduğunu öne sürmek için sıfır hipotezi testini ve güven aralıklarını kullanırlar. Tekrarlı ölçümler desenleri kullanan araştırmacıların, verilerin anlamını nasıl onayladıklarını göstermek için zaman algısı deneyini kullanacağız.

Zaman algısı deneyinin analizinde odak noktası, katılımcıların farklı uzunluklardaki zaman aralıklarını ayırt edip etmedikleridir. Tablo 7.3'teki ortalama farkların, hata değişkenliği açısından beklenenden daha büyük olduğu gösterilmedikçe, katılımcıların farklı uzunluktaki aralıkları ayırt edebildiklerini iddia edemeyiz. Bu durumda, katılımcılar farklı zaman aralıkları arasında ayrım yapabiliyor gibi görünse de bu farkın şans eseri ortaya çıkıp çıkmadığını bilemeyiz. Bu nedenle, bağımsız değişkenin etkili olup olmadığına karar vermemizde bize yardımcı olması için, sıfır hipotezini test etme ve güven aralıklarını oluşturma gibi analitik yöntemlerin kullanımını ele almalıyız.

Tekrarlı ölçümler desenlerinde yapılan analizin ayırt edici bir özelliği, hata değişkenliğinin tahmin edilme şeklidir. Bölüm 6'da, seçkisiz gruplar deseninde grup içi katılımcılar arasındaki bireysel farkların, hata değişkenliğinin tahmin edilmesini sağladığını belirttik. Böyle olmakla birlikte tekrarlı ölçümler desenlerinde katılımcılar arasındaki farklar sadece dengelenmemiştir; bu farklar

İLERİ ALIŞTIRMA

Bu alıştırmada, aşağıdaki tamamlanmış tekrarlı ölçümler deseninde, bağımsız değişkenin her düzeyi için ortalamayı hesaplamalısınız. İlk olarak üç koşula ait sonuçları özetlemeden ve betimlemeden önce, her koşulda yer alan her katılımcı için bir özet puan hesaplamalısınız.

Bir algı deneyinde, üç katılımcının karmaşık görsel örüntüleri belirleme yetenekleri test edildi. Her sunumda, katılımcılar kısa bir süre karmaşık bir örüntü (hedef) gördüler. Bu hedeften sonra dört seçenekli bir test verildi (hedefin kendisi ve hedefe benzeyen başka üç örüntü). Katılımcıların görevi, bu seçenekler arasından hedef örüntüyü seçmekti. Bağımsız değişken, sunum ile test arasında geçen 10, 30 ve 50 saniyelik zaman aralığıydı. Altı denemenin her birinde, katılımcılar bağımsız değişkenin bir düzeyinde 50 yargıda bulundular. Aşağıdaki Tablo her denemede her katılımcı için ABBA karşıt dengelenmiş deneme dizilerini göstermektedir. Parantez içindeki değerler, her denemede (maksimum 50) her katılımcının yaptığı hata sayısını (bağımlı değişken) göstermektedir. Bu tabloyu, gecikme süresinin (bağımsız değişkenin) hata sayısına olan etkisini betimlemek için kullanın.

Katılımcı	Deneme 1	Deneme 2	Deneme 3	Deneme 4	Deneme 5	Deneme 6
1	30s (9)	50s (6)	10s (2)	10s (6)	50s (10)	30s (3)
2	50s (10)	30s (6)	10s (2)	10s (4)	30s (8)	50s (8)
3	10s (1)	50s (6)	30s (7)	30s (3)	50s (8)	10s (3)

gerçekten analiz dışı bırakılmıştır. Tekrarlı ölçümler deseninin katılımcılara bağlı sistematik değişkenliği dışarda bırakma yeteneği, bu desenleri genellikle seçkisiz gruplar desenlerinden daha duyarlı yapar. Tekrarlı ölçümler desenlerindeki hata değişkenliğinin kaynağı, koşulların farklı katılımcıları farklı şekilde etkilemesidir.



Hata kaynağının tekrarlı ölçümler deseninde, bağımsız gruplar desenindekinden farklı tahmin edilmesi durumu, sıfır hipotezini test etmek için kullanılan *t* testi ve *F* testi hesaplamasının da farklı olmasına neden olur. Benzer şekilde, güven aralıkları da farklı hesaplanır. Bölüm 12'de tekrarlı ölçümler deseninin bir parçası olarak, hem *F* testinin hem de güven aralıklarının karar vermede nasıl kullanıldığını göstermek için Tablo 7.3'teki verileri kullanacağız. Tablo 7.3'teki verilerin analizi için sıfır hipotezi şudur: Örneklem ortalamalarına göre tahmin edilmiş evren ortalamaları, farklı zaman aralıkları için aynıdır. Bu verilerin varyans analizini yaparken (Bölüm 12'ye bakınız) diyebiliriz ki aralık uzunluğunun etkilerine ait *F* testi olasılığı $p = .0004$ idi. Bu değer .05'ten küçük olduğu için, zaman aralığı uzunluğunun etkisi istatistiksel olarak anlamlıdır. Bu sonucu temel alarak, katılımcıların zaman aralığı uzunluğuna dayalı zaman tahminlerinin, sistematik olarak farklılaştığını iddia edebiliriz. Hesaplamamız (eta kare = .80), etki büyüklüğünün ise büyük olduğunu zaten gösteriyor.

Bölüm 12'de .95 güven aralıklarını hesaplamak için yine Tablo 7.3'teki ortalamaları kullandık. Dört koşul için güven aralıkları (saniye cinsinden) (12) 5.4–19.8; (24) 14.8–29.2; (36) 30.8–45.2; (48) 32.2–46.6'dır. Bölüm 6'da (ayrıca Kutu 11.5'e bakınız) öğrenmiş olduğunuz gibi, zaman aralıkları farklı olduğunda, örneklem ortalamalarına dayanarak tahmin edilen evren ortalamalarının farklı olduğunu iddia edebiliriz. Bu zaman aralıklarının gözlenmesi, hangi ortalamaların farklı olduğuna karar vermenizi sağlar mı? Güven aralıkları arasındaki ilişkiyi incelemenin uygun bir yolu, grafik çizerek göstermektir. Örneğin, Bölüm 12'deki Şekil 12.2'ye baktığınız zaman göreceksiniz ki burada sunulan zaman aralıkları, zaman tahmini deneyinde elde edilen örneklem ortalamaları etrafında çizilmiştir.

AYRIMSAL AKTARMA SORUNU

- Bir koşula ilişkin etkiler sürekli ise ve sonraki koşulların performansını da etkiliyorsa, ayrımsal aktarım ortaya çıkar.
- Ayrımsal aktarıma yol açabilen değişkenler, seçkisiz gruplar desenleri kullanılarak test edilmelidir, çünkü ayrımsal aktarım tekrarlı ölçümler desenlerinin iç geçerliğini tehdit eder.
- Ayrımsal aktarım, bir tekrarlı ölçümler deseni ve bir seçkisiz gruplar deseninde test edildiğinde, aynı bağımsız değişkene ait sonuçlar karşılaştırılarak belirlenebilir.

Araştırmacılar alıştırma etkilerini dengelemek için kullanacakları uygun tekniklerle, tekrarlı ölçümler desenindeki potansiyel alıştırma etkisi probleminin üstesinden gelebilirler. Tekrarlı ölçümler desenlerinde gözlenen ve ayrımsal aktarım olarak bilinen çok daha ciddi bir potansiyel sorun vardır (Poulton, 1973, 1975, 1982; Poulton & Freeman, 1966). Bir koşuldaki performans, kendisinden önceki koşula dayalı olarak farklılaşırsa, **ayrımsal aktarım** ortaya çıkar.

Temel Kavram

Tekrarlı ölçümler deseninde iki farklı yönergenin karşılaştırıldığı bir problem çözme deneyini ele alın. Bir yönergeler kümesi (B) nötr kontrol koşulu temsil ederken, bir diğer yönergeler kümesinin (A) problem çözmeyi arttırması beklenmektedir. AB sırasıyla test edilen katılımcıların, B yönergelerini izlemeleri istendiğinde, A yönergeleri ile belirlenen sınırlı yaklaşımı bırakamamaları ya da bırakmakta isteksiz olmalarını beklemek akla yakın görünüyor. A yönergesini almış olan katılımcılar "İyi şey"i bırakarak, "Pembe filleri düşünme " yönergesini başarıyla uygulamış oluyorlar. Katılımcıların B yönergesini izlemeleri istenirken ilk koşuldaki (A) yönergesini terk etmeyi başaramadıkları zaman, iki koşul arasındaki fark azalır. Deneyin sonunda bu katılımcılar, gerçekte B koşulunda denenmemişlerdir. Bu deneyde katılımcıların "AB" koşulunda değil, "AA" koşulunda test edildiği bir durum karşımıza çıkar.

Genelde ayrımsal aktarım meydana gelmesi iç geçerliği tehdit eder, çünkü koşullar arasında gerçekten farklılıklar olup olmadığını belirlemek olanaksız hâle gelir. Ayrımsal aktarım aynı zamanda, koşullar arasındaki farklılıkları olduğundan düşük göstermeye eğilimlidir ve bu durum bulguların dış geçerliğini

düşürür. *Bu nedenle ayrımsal aktarımın ortaya çıkabildiği durumlarda, araştırmacılar bir bağımsız gruplar deseni seçmelidir.* **Yönerge değişkenlerinde ayrımsal aktarım, bu araştırmalarda tekrarlı ölçümler desenlerinin kullanılmamasını önermeye yetecek kadar yaygındır** (Underwood & Shaughnessy, 1975). Ne yazık ki ayrımsal aktarım herhangi bir tekrarlı ölçümler deseninde ortaya çıkabilir. Örneğin, araştırmacılar çeşitli sağlık problemlerine karşı gelişen damgalamayı araştırmak için kısa hikâyeler kullandıkları zaman, olası dizilim etkilerine karşı dikkatli olmalıdırlar. Bir araştırmada katılımcılar kanserli bir hastanın anlatıldığı kısa bir hikâyeye, deneysel düzende önce yer aldığı koşulda, sonra yer aldığı koşula kıyasla, farklı tepkide bulundular (Horch & Hodgins, 2008). Böyle olmakla birlikte ayrımsal aktarımın ortaya çıkma olasılığının olup olmadığını belirlemenin yolları vardır.

Bir tamamlanmamış tekrarlı ölçümler deseninde, ayrımsal aktarımın ortaya çıkıp çıkmadığını belirlemenin bir yolu, *sadece* ilk dizilimde yer alan sonuçları incelemektir. Örneğin, dört kısa hikâyenin verildiği damgalama deneyinde, katılımcıların dörtte biri önce kumar oynama koşulunda, diğer dörtte biri önce şizofreni koşulunda yer aldı ve bu şekilde devam etti. Bu nedenle, Tablo 7.2'de yer alan tüm olası ve seçilmiş sıralamalarda gösterildiği gibi ilk dizilim, dört gruplu bir bağımsız gruplar desenini temsil etmektedir. *Eğer yeterli sayıda katılımcı varsa,* araştırmacı bağımsız değişkenin etkisini bu dört grup için analiz edebilir ve daha sonra bu bulguları, tekrarlı ölçümler desenine ait tüm bulgularla karşılaştırabilir. Eğer bulgular farklılaşırsa, ayrımsal aktarım ortaya çıkmış olması olasıdır.

Ayrımsal aktarımın bir problem oluşturup oluşturmadığını belirlemenin en iyi yolu, iki ayrı deney yapmaktır (Poulton, 1982). Bu durumda aynı bağımsız değişken her iki deneyde de araştırılmış olacaktır, ama bir deneyde seçkisiz gruplar deseni, diğer deneyde ise tekrarlı ölçümler deseni kullanılmış olacaktır. Seçkisiz gruplar deseni ayrımsal aktarım içermeyebilir çünkü her katılımcı sadece bir koşulda test edilir. Eğer tekrarlı ölçümler deseni ile de seçkisiz gruplar deseninde gösterilmiş olan bağımsız değişken etkisi gösterilirse, ayrımsal aktarım olmamış olabilir. Ama iki desen aynı bağımsız değişken için farklı etkiler gösterirse, tekrarlı ölçümler deseninde farklı bir sonuç çıkmasından ayrımsal aktarım sorumlu olabilir. Ayrımsal aktarım ortaya çıktığı zaman, bağımsız değişkenin etkisini en iyi şekilde betimlemek için seçkisiz gruplar deseninin sonuçları kullanılmalıdır.

ÖZET

Tekrarlı ölçümler deseni, deneydeki her deneğe tüm koşulları uygulayarak deneyin yapılmasını sağlayan etkin ve yeterli bir yöntemdir (bk. Şekil 7.4). Tekrarlı ölçümler desenleri ancak araştırmada katılımcı sayısı çok az ise veya bağımsız değişken az sayıda katılımcıyı bir kaç kez test ederek daha etkin şekilde çalışılabiliyorsa faydalıdır. Tekrarlı ölçümler desenleri genellikle daha duyarlı deneylerdir. Nihayetinde, psikolojinin özel alanları (örneğin, psikofizik) tekrarlı ölçümler desenlerinin kullanımını gerektirir.

ŞEKİL 7.4 Bu bölümde tekrarlı ölçümler desenlerine ve karşıt dengeleme yöntemlerine giriş yaptık.



Bununla birlikte, herhangi bir tekrarlı ölçümler deseninin yorumlanabilir olması için, alıştırma etkileri dengelenmelidir. Alıştırma etkileri, testin tekrar edilmesi nedeniyle katılımcıların uğradıkları değişimlerdir. Bir tamamlanmış tekrarlı ölçümler deseninde, alıştırma etkileri her katılımcı için dengelenir. Blok seçkisizleştirme ve ABBA karşıt dengeleme, tamamlanmış tekrarlı ölçümler desenindeki alıştırma etkilerini dengelemede kullanılabilmektedir. Bununla birlikte, eğer beklenti etkisi veya alıştırma etkisinin doğrusal olmaması bekleniyorsa, ABBA karşıt dengeleme kullanılmamalıdır.

Bir tamamlanmamış tekrarlı ölçümler deseninde her katılımcı her uygulamayı sadece bir kez alır ve alıştırma etkileri katılımcılar aracılığıyla dengelenir. Tamamlanmamış tekrarlı ölçümler desenindeki alıştırma etkilerini dengelemede kullanılan teknikler, tüm olası ya da seçilmiş sıralamaların kullanımını içerir (Latin Karesi ve seçkisiz başlayan dönüşümlü dizilim yöntemi).

Tekrarlı ölçümler deseninde veri analizi süreci, temelde seçkisiz gruplar desenindekiyle aynıdır. Tamamlanmış tekrarlı ölçümler deseni için eklenen bir basamak, her katılımcının puanlarının, öncelikle her koşul içinde özetlenmesidir. Veriler, hatalar ayıklandıktan sonra ortalama, standart sapma ve etki büyüklüğü gibi betimleyici istatistikler kullanılarak özetlenir. Bağımsız değişkenin davranışı etkilediğini gösterebilmek için sıfır hipotezi testi ve güven aralıkları kullanılır.

Herhangi bir tekrarlı ölçümler desenindeki en ciddi sorun, ayrımsal aktarımdır –bir koşuldaki performansın, kendisinden önceki koşulun performansına bağlı olarak değişmesi durumu. Ayrımsal aktarımın varlığını belirlemede kullanılan yöntemler vardır ama ayrımsal aktarım meydana geldikten sonra, araştırmayı kurtarmak çok zordur.

TEMEL KAVRAMLAR

tekrarlı ölçümler desenleri 220
duyarlılık 221
alıştırma etkileri 223
karşıt dengeleme 224
ayrımsal aktarım 237

GÖZDEN GEÇİRME SORULARI

1 Bir seçkisiz gruplar deseninde ve bir tekrarlı ölçümler deseninde ne dengelenir? Açıklayın.
2 Araştırmacıların tekrarlı ölçümler desenini kullanmalarını açıklayan dört nedeni kısaca betimleyin.
3 Duyarlılığı tanımlayın ve tekrarlı ölçümler desenlerinin, seçkisiz gruplar desenine göre niçin sıklıkla daha duyarlı olduğunu açıklayın.
4 Tekrarlı ölçümler desenlerinde tamamlanmış ve tamamlanmamış deseni birbirinden ayırın.
5 Tamamlanmış bir desenin kullanıldığı bir tekrarlı ölçümler deneyinde, alıştırma etkilerini dengelemek için araştırmacının seçenekleri nelerdir?
6 Tamamlanmış desenin kullanıldığı bir tekrarlı ölçümler deneyindeki alıştırma etkilerini dengelemek için, hangi iki koşulda ABBA karşıt dengeleme kullanımına karşı çıkarsınız?
7 Tamamlanmış desenin kullanıldığı bir tekrarlı ölçümler deneyinde alıştırma etkilerini dengelemenin genel kuralını açıklayın.
8 Tamamlanmamış bir desenin kullanıldığı tekrarlı ölçümler deneyindeki alıştırma etkilerini dengelemek için araştırmacıların kullanabileceği teknikleri kısaca açıklayın. Bu tekniklerin hangilerinin tercih edildiğini belirtin ve nedenini açıklayın.
9 Tamamlanmış tekrarlı ölçümler deseninin kullanıldığı bir deneyde, verileri özetlemek için, neden ek bir ilk adım gerektiğini açıklayın.
10 Bir tekrarlı ölçümler deseninde ayrımsal aktarım ortaya çıkarsa, araştırmacıların bunu nasıl belirleyebileceklerini açıklayın.

DÜŞÜNDÜRÜCÜ SORULAR

1 Aşağıdaki problemler tekrarlı ölçümler desenlerinde alıştırma etkilerinin dengelenmesini gerektiren farklı durumları göstermektedir.
   A Bir bağımsız değişkeni olan, tamamlanmış bir desenin kullanıldığı bir tekrarlı ölçümler deneyini düşünün. Deneydeki bağımsız değişken, üç düzeyi olan görev zorluğudur (az, orta ve çok zor). Bu deney koşullarını uygularken, bağımsız değişkeni alıştırma etkileri açısından dengeleyecek şekilde bir dizilim yapmalısınız. Alıştırma etkilerini dengelemek için öncelikle blok seçkisizleştirme ve sonra ABBA karşıt dengelemeyi kullanmalısınız (Her bir koşul, hazırladığınız dizilimde iki kez yer almalıdır. Seçkisiz bloklamada iki seçkisiz dizinizi belirlemek için ekteki Tablo A.1 seçkisiz sayı tablosunun ilk sırasını kullanabilirsiniz.).
   B Tamamlanmamış bir desenin kullanıldığı bir tekrarlı ölçümler deneyi düşünün. Deneydeki bağımsız değişken, yazı tipinin boyutudur ve altı düzeyi vardır (7, 8, 9, 10, 11 ve 12 punto). Bu koşulları deneyin ilk altı katılımcısına nasıl uygulayacağınızı gösteren bir tablo oluşturun.
2 Bir reklam ajansında stajyer olarak çalışan bir öğrenciye, insanların yeni bir reklama ilişkin ilk izlenimlerini değerlendirme görevi verildi. Ajans, yeni bir araba modeli olan ticari ürünün dört fotoğrafını hazırladı. Araba potansiyel alıcılara arabanın özelliklerini vurgulayacak şekilde dört farklı açıdan fotoğraflandı (örneğin, kent dışındaki bir evin yanında, doğa manzaralı bir otoyolda seyahat sırasında vb. gibi). Stajyer, bir AVM'de alışveriş yapanların arasından seçkisiz olarak seçtiği insanlara dört fotoğrafı göstermeye karar verir. Seçilen kişilere her fotoğrafa bakmak için beş dolar verilmekte ve gösterilen fotoğrafların bu

arabanın tanıtımı açısından iyi bir reklam aracı olup olmayacağına karar vermeleri istenmektedir.Dört fotoğraftan her biri laptop ekranında 100 msn gösterilmekte ve katılımcı tercihini belirtmek için 1 ile 5 arasındaki sayılardan birisini işaretlemektedir (1 = hiç iyi değil 5 = bu araba için çok iyi reklam). Katılımcılardan her fotoğrafı gördükten sonra, ilk izlenimlerine ilişkin kararlarını olabildiğince hızlı bir şekilde, söylemeleri istenmiştir. Değerlendirmelere ek olarak, bu karar verme zamanı da ölçülmüştür.

A Farklı fotoğrafların etkisini incelemek için hangi desen kullanılmalıdır?

B Bu deneydeki koşulları kullanarak alıştırma etkilerini dengelemek için bir Latin karesi hazırlayın (Fotoğrafları A, B, C ve D şeklinde işaretleyin.).

C Stajyerin alıştırma etkilerini dengelemek için tüm olası dizilimleri kullanmaya ve her bir katılımcıyı 24 olası dizilimin her birine atamaya karar verdiğini varsayın. Sadece bu deneyde ilk sıraya konulacaklar (yani, her bir katılımcının göreceği ilk fotoğraf) üzerinde düşünün. 24 katılımcının hepsinin sadece ilk sıradaki fotoğraflara baktığı durumda, hangi deney deseni kullanılacaktır? Dört koşulun her birinde kaç katılımcı vardır?

D C'ye verdiğiniz cevabınızı da göz önüne aldığınızda, bu stajyer alıştırma etkilerini dengelemek için tüm olası dizilimleri kullandığı zaman, ayrımsal aktarım ortaya çıkıp çıkmadığını nasıl test edebilirdi?

3 Aşağıdaki tablo, tamamlanmamış desenin kullanıldığı bir tekrarlı ölçümler deneyinde, koşulların katılımcılara hangi sırada uygulandığını göstermektedir. Deneyin bağımsız değişkeni, çocuklar için geliştirilmiş bir elektronik bulmaca oyununun zorluk düzeyidir. Dört yaş çocuklarına bir tablet verildi ve kendilerinden, üç farklı zorluk düzeyinde saklanan şekilleri bulma oyunu oynamaları istendi. Zorluk düzeyleri, ekranda görünen resimlerin içerisine yerleştirilmiş şekillerin boyutlarına göre belirlendi: Aşırı küçük (AK), çok küçük (ÇK) ve küçük (K). Bağımlı değişken, üç zorluk düzeyinin her birinde bir çocuğun bulduğu şekillerin sayısıydı (maksimum = 10). Altı çocuk test edildi; tablodaki değerler her zorluk düzeyinde çocukların buldukları gizlenmiş şekillerin sayısını göstermektedir. Tabloyu, aşağıdaki soruları cevaplamak için kullanın.

Katılımcı	Koşulların sırası		
1	AK (2)	ÇK (9)	K (9)
2	ÇK (3)	K (5)	AK (7)
3	K (4)	AK (3)	ÇK (5)
4	AK (6)	K (10)	ÇK (8)
5	ÇK (7)	AK (8)	K (6)
6	K (8)	ÇK (4)	AK (4)

A Bu deneyde alıştırma etkilerini dengelemek için hangi teknik kullanıldı?

B Zorluk değişkeninin toplam etkisini betimlemek için kullanacağınız değerleri gösterin. Bu etkiyi sözel olarak da betimleyin ve bu betimlemenize, kullandığınız betimleyici istatistiği dâhil edin.

C Eğer zorluk değişkeninin etkisine ilişkin F testi değeri $p = .04$ ise, zorluk değişkeninin etkisi hakkındaki iddianız ne olur?

İleri Alıştırmaların Cevapları

İlk adım, aynı koşula ait iki denemenin ortalamasını alarak, bağımsız değişkenin her bir düzeyi için her bir katılımcının ortalamasını hesaplamaktır. İlk katılımcının üç koşuldaki ortalamaları şöyledir:

10s	30s	50s
(2 + 6)/2 = 4	(9 + 3)/2 = 6	(6 + 10)/2 = 8

İkinci katılımcının üç koşuldaki ortalamaları 3 (10 s), 7 (30 s) ve 9 (50 s), ve üçüncü katılımcının ortalamaları ise 2 (10 s), 5 (30 s), ve 7'dir (50 s).

Sonraki adım, her katılımcı için özet puanların ortalamasını alarak her koşul için ortalamaları hesaplamaktır.

10s: (4 + 3 + 2)/3 = 3

30s: (6 + 7 + 5)/3 = 6

50s: (8 + 9 + 7)/3 = 8

Şimdi bağımsız değişkenin (hedef ile test arasındaki gecikme) bağımlı değişken (hata sayısı) üzerindeki etkisini betimleyebiliriz. Ortalamalar gösteriyor ki hedef ile test arasındaki gecikme arttıkça, örüntü tanıma görevindeki hata sayısı da artmıştır. Gecikme değişkeninin güvenilir bir etkisinin olup olmadığını doğrulamak için, çıkarımsal istatistiklerden sıfır hipotezi veya güven aralığı kullanılabilir.

Düşündürücü Soruların Cevapları

A 1, 2 ve 3 değerlerini sırasıyla Düşük, Orta ve Yüksek koşullarına tahsis eder ve ekteki seçkisiz sayı tablosunun (Tablo A.1) ilk sırasındaki ilk sayı ile başlarsak blok seçkisizleştirme dizilimi Düşük-Yüksek-Orta-Düşük-Orta-Yüksek şeklinde ortaya çıkar. Muhtemel bir ABBA karşıt dengeleme dizisi ise Düşük-Orta-Yüksek-Yüksek-Orta-Düşük şeklinde olabilir..

B Altı koşul olduğu için, tüm olası diziler uygundur. Bu yüzden, alıştırma etkilerini dengelemek için bir Latin Karesi veya seçkisiz başlayan dönüşümlü dizilim ile başlamak gerekir. Dönüşümü kullanan bir olası diziler seti şöyle olur:

	Sıra					
Katılımcı	1'inci	2'nci	3'üncü	4'üncü	5'inci	6'ıncı
1	8	10	11	9	7	12
2	10	11	9	7	12	8
3	11	9	7	12	8	10
4	9	7	12	8	10	11
5	7	12	8	10	11	9
6	12	8	10	11	9	7

BÖLÜM SEKİZ

Karmaşık Desenler

BÖLÜMÜN ANA HATLARI

GENEL BAKIŞ
KARMAŞIK BİR DESENDEKİ ETKİLERİN BETİMLENMESİ
Bir 2 x 2 Desen Örneği
Temel Etki ve Etkileşim
Etkileşimin Açıklanması
Üç Bağımsız Değişkenli Karmaşık Desenler
KARMAŞIK DESENLERİN ANALİZİ
Etkileşimin Olduğu Bir Durumda Analiz Planı
Etkileşimin Olmadığı Bir Durumda Analiz Planı
ETKİLEŞİM ETKİLERİNİ YORUMLAMAK
Etkileşim Etkileri ve Teori Test Etme
Etkileşim ve Dış Geçerlik
Etkileşim Etkileri ve Tavan ve Taban Etkileri
Etkileşim ve Doğal Gruplar Deseni
ÖZET

GENEL BAKIŞ

Altıncı ve yedinci bölümlerde araştırmacıların bir bağımsız değişkenin etkisini araştırmak için kullandıkları temel deneysel desenlere odaklandık. Bir bağımsız değişkenin her koşulda ayrı bir grup katılımcıya (bağımsız gruplar desenleri) ya da her koşulun tüm katılımcılara (tekrarlı ölçümler desenleri) nasıl uygulanabileceğini açıkladık. Deneysel araştırmanın temelleri üzerine yoğunlaşmak istediğimiz için tartışmamızı yalnızca bir bağımsız değişken içeren deneylerle sınırladık. Ancak tek bağımsız değişken içeren deneyler, çağdaş psikoloji araştırmalarındaki en yaygın deney türü değildir. Bunun yerine, araştırmacılar sıklıkla tek bir deneyde iki ya da daha fazla bağımsız değişkenin eş zamanlı olarak ele alındığı **karmaşık desenler** kullanırlar.

Temel Kavram

Karmaşık desenler, bağımsız değişkenlerin faktöryel birleşimini (combination) içerdiği için aynı zamanda *faktöryel desenler* olarak da adlandırılır. *Faktöryel kombinasyon* bir bağımsız değişkenin her düzeyinin, ikinci bir bağımsız değişkenin her düzeyi ile eşleştirilmesini içerir. Bu eşleştirme, her bir bağımsız değişkenin tek başına etkisinin (*temel etki*) ve bağımsız değişkenlerin birlikte ortaya çıkardığı etkinin (*etkileşim etkisi*) belirlenmesine olanak sağlar.

Karmaşık desenler bu noktada biraz karmaşık görünebilir ama bu bölüm boyunca ilerledikçe kavramların daha anlaşılır hâle geldiğini göreceksiniz. Karmaşık bir desendeki bağımsız değişkenleri incelemek için kullanılan deneysel desenlerin özelliklerini değerlendirerek başlayabiliriz. Daha sonra, temel etkilerin ve etkileşimin oluşturulması, analiz edilmesi ve yorumlanması için izlenecek işlemleri açıklayacağız. Karmaşık desenler için kullanılan analiz planları ile başlayacağız. Karmaşık desenlerde etkileşimin yorumlanmasına özel önem vererek bölümü bitireceğiz.

KARMAŞIK BİR DESENDEKİ ETKİLERİN BETİMLENMESİ

- Araştırmacılar bir deneyde iki veya daha fazla bağımsız değişkenin etkilerini incelemek için karmaşık desenler kullanırlar.
- Karmaşık desenlerde her bağımsız değişken bir bağımsız gruplar deseni ile ya da tekrarlı ölçümler deseni ile çalışılabilir.
- En basit karmaşık desen, her birisinin iki düzeyi olan iki bağımsız değişkenden oluşan 2 x 2 desenidir.
- Karmaşık bir desendeki farklı koşulların sayısı, her bir bağımsız değişkene ait düzey sayısı çarpılarak belirlenir (örneğin, 2 x 2 = 4).
- Daha güçlü ve etkin karmaşık desenler, bir bağımsız değişkene daha fazla düzey veya desene daha fazla bağımsız değişken dâhil edilerek oluşturulur.

Karmaşık bir desende, doğası gereği, birden fazla bağımsız değişken vardır. Karmaşık bir desendeki her bir bağımsız değişken, 6'ncı ve 7'nci bölümlerde açıklanan işlemler doğrultusunda bir bağımsız gruplar deseni ya da tekrarlı ölçümler deseni kullanılarak uygulanmalıdır.

Karmaşık bir desende hem bağımsız gruplar değişkeni hem de tekrarlı ölçümler değişkeni varsa, bu desene *karma desen* (mixed design) denir.

Mümkün en basit deney, iki düzeyde değişimlenen bir bağımsız değişken gerektirir. Aynı şekilde, mümkün en basit karmaşık desen de her birinin iki düzeyi olan, iki bağımsız değişken içerir. Karmaşık desenler, deneydeki her bağımsız değişkene ilişkin düzeylerin sayısının spesifik hale getirilmesiyle belirlenir. Bu durumda, 2 x 2'lik bir desen ("iki çarpı iki" şeklinde okunur) en temel karmaşık desendir. Teorik olarak, sınırsız sayıda karmaşık desen oluşturmak mümkündür. Çünkü çok sayıda bağımsız değişkenin etkisi araştırılabilir ve her bir bağımsız değişkenin farklı sayıda düzeyleri olabilir. Bununla birlikte, uygulamada dört ya da beş bağımsız değişken içeren deneylere daha seyrek rastlanırken, iki ya da üç bağımsız değişkenin dâhil edildiği çalışmalarla daha sık karşılaşılır. Bağımsız değişkenlerin sayısından bağımsız olarak, karmaşık bir desendeki koşulların sayısı, bağımsız değişkenlerin düzeylerinin çarpılmasıyla belirlenir. Örneğin, her birinin iki düzeyi olan iki bağımsız değişken varsa (2 x 2'lik bir desen), dört koşul vardır. 3 x 3'lük bir desende her birinin üç düzeyi olan iki bağımsız değişken vardır; yani bu durumda dokuz koşul vardır. 3 x 4 x 2'lik bir desende sırasıyla, her birinin üç, dört ve iki düzeyi olan üç bağımsız değişken ve toplamda yirmi dört koşul vardır. Tüm karmaşık desenlerin en başta gelen avantajı, bağımsız değişkenler arasındaki etkileşimleri belirleme fırsatı sağlamasıdır.

2x2'lik bir deseni anlamak, karmaşık desenleri anlamak için bir zemin oluşturur. Böyle olmakla birlikte, 2 x 2'lik bir desen, karmaşık desenlerin potansiyelini anlamanın ancak başlangıcıdır. Karmaşık desenler iki yoldan herhangi biri kullanılarak 2 x 2'lik desenin ötesine genişletilebilir. Araştırmacılar desendeki bağımsız değişkenlerin birisine veya her ikisine birden yeni düzeyler ekleyebilir; örneğin, böylece 3 x 2, 3 x 3, 4 x 2, 4 x 3 ve benzeri desenler üretebilir. Ayrıca, araştırmacılar 2 x 2'lik bir deseni, aynı deneyde bağımsız değişkenlerin sayısını artırarak geliştirebilirler. Her bir değişkenin düzeylerinin sayısı, iki ile daha büyük bir üst limit arasında yer alabilir. Üçüncü ya da dördüncü bir bağımsız değişkenin eklenmesiyle 2 x 2 x 2, 3 x 3 x 3, 2 x 2 x 4, 2 x 3 x 3 x 2 ve benzeri desenler elde edilebilir.

İlk olarak, 2 x 2'lik bir desen örneği üzerinde çalışarak, karmaşık bir desendeki temel etkileri ve etkileşimi göstereceğiz.

Bir 2 x 2 Desen Örneği

Temel etki ve etkileşimin niteliği, esasen tüm karmaşık desenlerde aynıdır, ancak en kolay şekilde 2 x 2'lik bir desende görülebilir. Bu desene bir örnek vermek için psikoloji ve hukuk alanındaki zengin literatürden yararlanacağız. Hukukî düzlemde sosyal bilimcilerin dokunmadıkları birkaç alan vardır. Jüri seçimi, görgü tanıklığının niteliği ve güvenilirliği, sanığın ırkı, jürinin verdiği kararlar ve avukatların savunmaları, araştırmacıların inceledikleri birçok konudan sadece bazılarıdır. Araştırmacıların, sorgulama için getirilen şüphelilerin yanlış itiraflarına yol açabilecek değişkenleri inceledikleri bir çalışmayı ele alacağız.

Kassin, Goldstein ve Savitsky (2003) bir şüpheliyi sorgulayan kişilerin, şüphelinin suçlu ya da suçsuz olduğuna dair beklentilerinin, kullandıkları sorgulama taktiklerini etkileyip etkilemediğini araştırmak için 2 x 2'lik bir desen kullandı. Araştırmacılar, suçsuz insanların yanlış itiraflarına yol açan etkenleri belirlemek için birçok çalışma yapmışlardır. Bu araştırmada araştırmacılar, yanlış itiraflara yol açabilecek olası nedenlerden birisinin, sorgulayıcıların *doğrulama yanlılığı* (confirmation bias) olabileceğini varsaymışlardır. Bu varsayım doğrultusunda, şüphelinin suçlu olduğuna dair sorgulayıcıların başlangıçtaki inançlarının, şüphelileri daha saldırgan bir şekilde sorgulamalarına, suçlu olduklarını varsayan bir tavırla sorular sormalarına ve şüphelinin savunmacı davranışlar sergilemesine (tabi bu durum şüphelinin suçlu olduğu şeklinde yorumlanır) yol açtığı varsayılmıştır. Genel olarak, davranışsal doğrulama teorisi üç bölümden oluşur: (1) algılayan kişi, hedef kişi hakkında bir inanç oluşturur; (2) algılayan kişi, hedef kişiye inancı ile tutarlı bir şekilde davranır ve (3) hedef kişi daha sonra algılayanın inancını destekleyecek şekilde tepki verir. Nihayetinde, adlî ceza bağlamındaki bu süreç, suçsuz bir insanın suçlu olduğunu itiraf etmesiyle sonlanabilir.

Kassin ve arkadaşları (2003), davranışsal doğrulama teorisini üniversite öğrencilerinden oluşan bir örneklemde, zekice tasarlanmış bir deneyle test ettiler. Öğrenci çiftleri deneye "sorgulayıcı" ve "şüpheli" olarak katıldılar. Sorgulayıcılardan, kilitli bir dolaptan 100 $'ın çalındığı bir vakayı çözmeye çalışan bir dedektif rolünü oynamaları istendi. "Şüpheli" rolünü oynayan öğrenciler ise, sorgu öncesinde hırsızlığın gerçekleştiği olay yerini onlara hatırlatan bir dizi talimat aldılar. Tüm şüphelilere, sorgulayıcıyı suçsuz olduklarına dair ikna etmeleri ve hırsızlık ile ilişkili itirafta bulunmamaları talimatı verildi. Sorgulayıcılara, hem bir itiraf elde etmek hem de şüphelinin gerçek suçlu olup olmadığını belirlemek gibi iki karşıt görev verildi.

İlk Bağımsız Değişken: Şüphelinin Durumu Gerçek bir sorguda şüphelinin davranışı, suçu gerçekten işleyip işlemediğine bağlı olarak değişir; bu nedenle araştırmacılar *şüphelinin durumu* bağımsız değişkenini, öğrencilerin gerçek *suçlu* olma veya olmama şeklinde iki düzeyde değişimlediler. *Suçlu* olma koşulunda, öğrencilere bir odaya girerek, bir cihazın arkasına gizlenmiş bir anahtarı bulmaları, ardından o anahtarla kilitli kabini açıp içinden 100 $'ı almaları, kabini tekrar kilitledikten sonra odadan ayrılmaları talimatı verildi. *Suçsuz olma* koşulunda, öğrencilerden aynı odaya yaklaşarak kapıyı çalmaları, içeriden gelecek cevabı beklemeleri (ki kapı hiç açılmadı) ve ardından da araştırmacı ile buluşmaları istendi. Öğrenci şüphelilerin yarısı suçlu rolüne, diğer yarısı ise suçsuz rolüne seçkisiz olarak atandı.

İkinci Bağımsız Değişken: Sorgulayıcının Beklentisi Sorgulamalar arasında doğrulama yanlılığı yaratmak için araştırmacılar *sorgulayıcı beklentisi* bağımsız değişkenini iki düzeyde değişimlediler. Araştırmacı, öğrenci sorgulayıcıların yarısını *suçlu beklentisi* koşuluna seçkisiz olarak atadı ve kendilerine, beş şüpheliden dördünün gerçekten suçu işlediğini söyledi. Bu koşuldaki sorgulayıcılar, sorguladıkları kişinin suçlu olması olasılığının yüksek (%80) olduğuna inandırıldı. *Suçsuz beklentisi* koşulunda ise sorgulayıcılar sorguladıkları kişinin suçlu olması olasılığının düşük (%20) olduğuna inandırıldı, çünkü sorgulanacak beş şüpheliden yalnızca birinin hırsızlık suçunu işlediği söylendi.

İki Bağımsız Değişkenin Faktöryel Kombinasyonu Bu 2 × 2'lik karmaşık desende iki bağımsız değişkenin faktöryel kombinasyonuyla dört koşul yaratılmıştır:

1 Gerçek suçlu / suçlu beklentisi
2 Gerçek suçlu / suçsuz beklentisi
3 Gerçek suçsuz / suçlu beklentisi
4 Gerçek suçsuz / suçsuz beklentisi

Değişkenlerin kombinasyonuyla oluşturulan her grubun, rastlantısal bir grup katılımcıyı temsil ettiğini unutmayın. Desen aşağıdaki gibidir:

	Sorgulayıcı Beklentisi	
Şüpheli Statüsü	Suçlu	Suçsuz
Gerçek suçluluk	1	2
Gerçek suçsuzluk	3	4

Araştırmacılar davranışsal doğrulama teorisini test etmek için bazı bağımlı değişkenleri ölçmüşlerdir. Örneğin, sorgulayıcılar, şüpheliler ve kasete kaydedilmiş sorgulamaları dinleyen yeni, ek katılımcılar (daha çok olası jüri üyeleri gibi) için bağımlı değişken ölçümleri alınmıştır. Temel etkileri ve etkileşimleri örneklemek için deneydeki üç bağımlı değişken üzerinde duracağız. Şimdi, ne bulduklarını görelim.

Temel Etki ve Etkileşim

- Karmaşık bir desendeki her bir bağımsız değişkenin genel etkisine temel etki denir; bu etki, diğer bağımsız değişkenin düzeylerini hesaba katmadan (onları yok sayarak) bir bağımsız değişkenin her bir düzeyinin ortalamaları arasındaki farkları temsil eder.
- Bir bağımsız değişkenin etkisi, ikinci bir bağımsız değişkenin düzeylerine bağlı olarak değiştiği zaman, bağımsız değişkenler arasında etkileşim ortaya çıkar.

Temel Kavram

Herhangi bir karmaşık desende, diğer bağımsız değişken(ler)in etkisini göz ardı ederek, deneydeki her bir bağımsız değişkenin genel etkisi ile ilgili varsayımları test etmek mümkündür. Karmaşık bir desende bir bağımsız değişkenin genel etkisine **temel etki** denir. Kassin ve arkadaşlarının iki farklı bağımlı değişken için deneylerinde gözledikleri iki temel etkiyi inceleyeceğiz

Bağımlı Değişken #1: Suç-Varsayımına Dayalı Soruların Sayısı Sorgulamadan önce, sorgulayıcı öğrencilere hırsızlık hakkında sorabilecekleri olası sorulardan oluşan bir listeyi de içerecek şekilde sorgulama teknikleri hakkında bilgi verildi. On iki soru çiftler hâlinde yazıldı (ama listede seçkisiz şekilde sunuldu). Soru çiftlerinin birisi şüphelinin suçlu olduğunu varsayarken (örneğin, "Cihazın arkasında gizlenen anahtarı nasıl buldunuz?") ikinci soruda böyle bir varsayım yoktu (örneğin, "Cihazın arkasında gizlenen anahtar hakkında bir şey biliyor musunuz?"). Sorgulayıcı öğrencilerden daha sonra

TABLO 8.1 SORGULAYICI BEKLENTİSİNİN SUÇ VARSAYIMINA DAYALI SORU SAYISI ÜZERİNDEKİ TEMEL ETKİSİ

	Sorgulayıcı Beklentisi	
Şüpheli Statüsü	Suçlu	Suçsuz
Gerçek suçluluk	3.54	2.54
Gerçek suçsuzluk	3.70	2.66
Sorgulayıcı beklentisinin ortalaması	3.62	2.60

Kassin ve ark. (2003) çalışmasına dayanan hipotetik hücre ortalamaları

sormak isteyebilecekleri altı soru seçmeleri istendi; öğrenciler, suç varsayan 0'dan 6'ya kadar olan sorular arasından soru seçtiler. Davranışsal doğrulama teorisine göre, suçlu beklentisi koşulunda yer alan sorgulayıcıların, suçsuz beklentisi koşulundaki sorgulayıcılara kıyasla, daha fazla sayıda suç varsayımına dayalı soru seçmeleri beklenir. Böylece araştırmacılar sorgulayıcı beklentisi bağımsız değişkeninin *temel etkisini* yordadılar.

Suç varsayımına dayalı soru sayısı bağımlı değişkenine ait veriler Tablo 8.1'dedir. Suçlu beklentisi koşulunda katılımcılar için suç varsayımına dayalı soruların sayısının genel ortalaması (3.62), gerçek suçluluk (3.54) ile gerçek suçsuzluk (3.70) koşullarının ortalaması alınarak bulunur. Benzer şekilde, suçsuz beklentisi koşulu için genel ortalama (2.54 + 2.66) / 2 = 2.60 olarak hesaplanır.[1] Bir temel etkinin ortalamaları, diğer bağımsız değişkenin düzeylerini hesaba katmadan, belirli bir bağımsız değişkenin her bir düzeyindeki performansı temsil eder. Örneğimizde sorgulayıcı beklentisi değişkeninin temel etkisinin ortalamasını elde etmek için, şüpheli statüsü değişkeninin ortalamasını aldık. Sorgulayıcı beklentisi değişkeninin temel etkisi, suç varsayımına dayanan seçilmiş soruların tümünün, sorgulayıcıların suçlu beklentisi içinde oldukları durumda (3.62), olmadıkları duruma (2.60) kıyasla daha büyük olduğunu gösterir. Çıkarımsal istatistik testleri, sorgulayıcı beklentisinin temel etkisinin istatistiksel olarak anlamlı olduğunu doğruladı. Bu sonuç ise araştırmacıların davranış doğrulama teorisine dayalı hipotezlerini destekledi.

Bağımlı Değişken #2: İkna Edici Tekniklerin Sayısı Şimdi şüpheli statüsü bağımsız değişkeni temel etkisinin, istatistiksel olarak anlamlı olduğu bir bağımlı değişkene dönelim. Araştırmacılar aynı zamanda sorgulayıcıların itiraf elde etmek için kullandıkları teknikleri analiz etmek için görüşmeleri videoya kaydettiler. Öğrenci sorgulayıcılara, şüphelinin direncini kırmak için polisin kullandığı güçlü teknikler hakkında kısa ve yazılı yönergeler verildi.

[1]Temel etkilere ilişkin ortalamaları elde etmek için, her sıra ve sütun içindeki değerlerin basitçe ortalamasının alınması, sadece tablodaki her ortalamaya katkı sağlayan eşit sayıda katılımcı olduğu zaman mümkündür. Katılımcı sayısı farklı olduğunda, ağırlıklı ortalamaların nasıl hesaplanacağına ilişkin olarak Keppel'e (1991) bakın.

TABLO 8.2 ŞÜPHELİ STATÜSÜNÜN İKNA EDİCİ TEKNİKLERİN SAYISI ÜZERİNDEKİ TEMEL ETKİSİ

Şüpheli Statüsü	Sorgulayıcı Beklentisi		Şüpheli Statüsü Ortalamaları
	Suçlu	Suçsuz	
Gerçek suçluluk	7.71	6.59	7.15
Gerçek suçsuzluk	11.96	10.88	11.42

Kassin ve arkadaşlarının (2003) çalışmasına dayanan hipotetik hücre ortalamaları

Araştırmacılar yakınlık kurma, şüphelinin suçlu olduğu iddiası ya da şüphelinin ifadelerine inanmama, şüphelinin ilgisini çekme ya da vicdanını harekete geçirme, ceza tehdidi, hoşgörü gösterileceğine dair vaat ve sahte kanıt sunulması gibi ikna amacı güden sorgulayıcı ifadelerin sayısını belirlediler.

Bu bağımlı değişkene, yani ikna edici teknik sayısına ait veriler. Tablo 8.2'dedir. Gerçek suçlu olan şüphelilerle görüşürken sorgulayıcıların kullandıkları ikna edici tekniklerin ortalaması 7.15 idi. Bu ortalama, gerçekte suçlu koşulunda sorgulayıcı beklentisi değişkeninin iki düzeyinin ortalaması alınarak hesaplanmaktadır: (7.71 + 6.59)/2. Şüphelinin gerçekte suçsuz olduğu koşulda, sorgulayıcıların kullandıkları tüm ikna edici tekniklerin ortalaması, 11.42 idi; bu rakam, sorgulayıcı beklentisi değişkeninin gerçekte suçsuz koşulundaki ortalaması alınarak hesaplandı: (11.96 + 10.88)/2. Ortalama olarak gerçekten *suçsuz* şüpheliler, suçlu olanlarla karşılaştırıldığında, sorgulayıcılar 4.27 daha fazla ikna edici teknik kullanmışlardır (11.42 − 7.15 = 4.27). Sorgulayıcı beklenti koşullarının ikisinde de suçsuz şüpheliler, gerçekten suçlu olan şüphelilerden daha fazla saldırgan bir şekilde sorgulanmışlardır.

Bağımlı Değişken #3: İtiraf Ettirme Çabası Son olarak, araştırmacıların aynı zamanda şüphelinin statüsü ile sorgulayıcının beklentisi arasındaki etkileşimi gözledikleri verileri de inceleyebiliriz. Deneyin ikinci aşamasında yeni bir öğrenci grubu, teybe kaydedilmiş bir sorgulama videosu dinlediler ve sorgulayıcının ve şüphelinin davranışları hakkında hüküm verdiler. Sorgulayıcının şüpheliden itiraf almak için gösterdiği çabanın derecesini belirlemek üzere, öğrencilere 10 dereceli bir ölçek verildi; bu ölçekte rakam büyüdükçe gösterilen çaba da büyüyordu. Bu veriler Tablo 8.3'tedir.

TABLO 8.3 SORGULAYICI BEKLENTİSİ VE ŞÜPHELİ STATÜSÜ ARASINDAKİ ETKİLEŞİM ETKİSİNİN İTİRAF ALMA ÇABASI ÜZERİNDEKİ ETKİSİ

Şüpheli Statüsü	Sorgulayıcı Beklentisi	
	Suçlu	Suçsuz
Gerçek suçluluk	5.64	5.56
Gerçek suçsuzluk	7.17	5.85

Hücre ortalamaları Dr. Saul Kassin'dan alınmıştır.

Temel Kavram

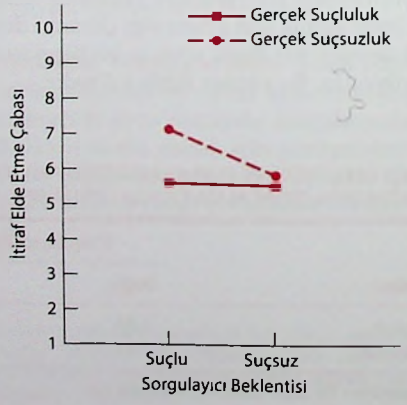
İki bağımsız değişken arasında etkileşim olduğu zaman biliyoruz ki her iki değişken, bağımlı değişken puanlarını (örneğimizde, sorgulayıcıların itiraf ettirmek için sarf ettikleri gayretin derecesini) birlikte etkiliyor. Formel olarak ifade edersek bir **etkileşim etkisi,** bir bağımsız değişkenin etkisi ikinci bir bağımsız değişkenin düzeyine bağlı olarak değiştiği zaman ortaya çıkar. Etkileşimi daha iyi anlayabilmek için Tablo 8.3'te ilk satırı inceleyin. Eğer deneyde gerçekten suçlu şüpheliler ile çalışılmış olsaydı, sorgulayıcı beklentisinin, sorgulayıcının çabasına ilişkin değerlendirme üzerinde *etkisinin olmadığı* sonucuna varacaktık; çünkü suçlu ve suçsuz beklentisi koşulları için ortalamalar neredeyse aynıdır. Öte yandan, gerçekten suçsuz olan şüpheliler test edilmiş olsaydı, (ikinci satır), sorgulayıcı beklentisinin itiraf elde etme çabası üzerinde *büyük bir etkisinin* olduğuna karar verecektik.

Etkileşim, koşullara ait ortalamaların grafiği çizildiği zaman daha kolay görülür. Tablo 8.3'te yer alan dört ortalama Şekil 8.1'de görülmektedir. Bu sonuçlar, sorgulayıcıların çabasına ilişkin değerlendirmelerin, şüphelinin gerçekten suçsuz ya da suçlu olması ve sorgulayıcının şüphelinin suçlu ya da suçsuz olduğuna dair beklentisine dayandığını gösterir. Bu durum, etkinin olup olmadığını açıklamak için iki bağımsız değişkenin de gerekli olduğunun göstergesidir. Bir sonraki bölümde "Karmaşık Desenlerin Analizi" başlığı altında karmaşık desenlerde etkileşimin analizini istatistiksel olarak açıklayacağız. Şimdilik *etkileşimin, bir bağımsız değişkenin etkisinin, ikinci bir bağımsız değişkenin düzeylerine bağlı olarak değiştiği durumda* meydana geldiğini bilmemiz yeterlidir.

Bir bağımsız değişken ikinci bir bağımsız değişkenle etkileşime girdiğinde, ikinci bağımsız değişken birincisi ile etkileşime girmelidir (Yani, bağımsız değişkenlerin sırası önemli değildir.). Örneğin, Tablo 8.3'te sorgulayıcıların beklentisine ilişkin etkinin şüphelilerin statüsüne bağlı olduğunu göstererek etkileşimi açıkladık. Tersi de doğrudur; şüpheli statüsünün etkisi de sorgulayıcıların beklentisine bağlıdır.

Şimdi biz, Kassin ve arkadaşlarının (2003) tüm veri analizlerine dayanarak vardıkları sonuçları açıklayabilecek durumdayız. Davranış doğrulama teorisini kullanarak araştırmacılar, sorgulayıcıların suça yönelik beklentilerinin, kendi

ŞEKİL 8.1 İtiraf elde etme çabası üzerine sorgulayıcı beklentisi ve şüpheli statüsü arasındaki etkileşim etkisinin grafiksel gösterimi (Dr. Saul Kassin tarafından sağlanmış veridir)



İLERİ ALIŞTIRMA I

Bu alıştırmada Tablo 8.1, 8.2 ve 8.3'ü inceleyerek aşağıdaki soruları cevaplamanız gerekiyor.

1 (a) Tablo 8.1'de şüpheli statüsü bağımsız değişkeninin temel etkisine ait ortalamalar hangileridir?
(b) Bu veriler temel alındığında, şüpheli statüsü değişkeninin temel etkisi, sorgulayıcı beklentisi değişkeninin temel etkisi ile nasıl karşılaştırılabilir?
c) Bu verilerde bir etkileşim etkisi olabilir mi?

2 (a) Tablo 8.2'de sorgulayıcı beklentisi bağımsız değişkeninin temel etkisine ait ortalamalar hangileridir?
(b) Bu verilere göre, sorgulayıcı beklentisi değişkeninin temel etkisi, şüpheli beklentisi değişkeninin temel etkisi ile nasıl karşılaştırılabilir?
c) Bu verilere göre, bir etkileşim etkisi olabilir mi?

3 (a) Tablo 8.3'te sorgulayıcı beklentisi bağımsız değişkeninin temel etkisine ait ortalamalar hangileridir?
(b) Tablo 8.3'te şüpheli statüsü bağımsız değişkeninin temel etkisine ait ortalamalar hangileridir?
(c) Kassin ve arkadaşları (2003) bu temel etkilerin istatistiksel olarak anlamlı olduğunu gözlemlediler. Hesapladığınız ortalamaları kullanarak, Tablo 8.3'teki sorgulayıcı beklentisi ve şüpheli statüsü değişkenlerinin temel etkilerini betimleyin.

inançlarını doğrulayacak bir sorgulama yürütmesine neden olacağını öngördüler. Genel olarak araştırmanın sonuçları hipotezlerini desteklemiştir; şüphelinin suçlu olduğundan şüphe eden sorgulayıcılar, daha saldırgan görüşmeler yürütmüşlerdir. Bu koşuldaki şüpheliler ise daha savunmacı olmaya başlamış ve tarafsız gözlemciler de bunları suçlu olarak algılamışlardır. Suçluluk beklentisi içindeki sorgulayıcıların, aslında suçsuz olan şüphelilerden itiraf elde etmeye çalışırken bile, daha saldırgan olmaları, suç beklentisinin ve davranışsal doğrulama sürecinin gücünü gösterir. Adalet sistemi bağlamında şüphelinin suçlu olduğuna dair ön yargıya dayanan polis sorgulamaları ve suçsuz insanlardan elde edilen sahte itiraflar trajik sonuçlara yol açabilen bir dizi yanlılığı tetikleyebilir.

Etkileşimin Açıklanması

- Etkileşimin kanıtları grafiklerle (örneğin, paralel olmayan doğrularla) veya tablolarla (çıkarma yöntemi) gösterilen betimleyici istatistiklerle gösterilebilir.
- Etkileşimin varlığı, çıkarımsal istatistikler kullanılarak doğrulanır.

Etkileşimin sonuçlarını betimleme şeklini nasıl seçeceğiniz, bu etkiyi hangi açıdan vurgulamak istediğinize bağlıdır. Örneğin, Kassin ve arkadaşları (2003) davranışsal doğrulama teorisine dayanan kestirimlerini test etmek için sorgulama beklentisi değişkeninin suçsuz ve suçlu şüpheliler üzerindeki etkisini ön planda tuttular. Yani sorgulayıcıların, şüphelinin suçlu ya da suçsuz olduğuna dair beklentisini değişimleyerek, "sorgulayıcıların bu beklentilerine uygun kanıt bulmaya çalışacakları" kestirimini test ettiler. İkinci bağımsız değişken (gerçekte suçlu veya suçsuz olmak) eklenerek, araştırma gerçek hayattaki sorgulamalara daha fazla benzetildi. İkincisi, araştırmacılar şüphelilerin suçsuzluk iddialarına rağmen suçlu beklentisi olan sorgulayıcıların, itiraf ettirmek için daha fazla uğraştığını gösterdi. Karmaşık desenlerde etkileşim etkisinin araştırılması, sadece bir bağımsız değişkenle yapılan deneylere kıyasla, araştırmacılara daha fazla bilgi edinme olanağı sağlar.

İLERİ ALIŞTIRMA II

Bu alıştırmada 2 × 3 karmaşık desenlerde sadece betimsel istatistik kullanarak, temel etkileri ve etkileşim etkilerini belirlemek için alıştırma yapma olanağınız olacak.

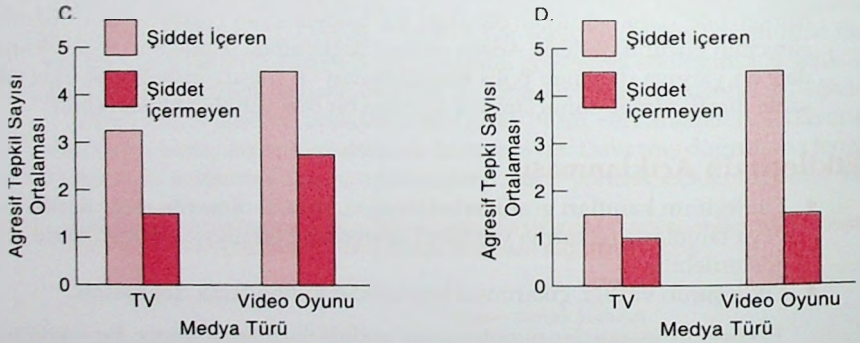
Göreviniz, altı karmaşık desen deneyinin (A'dan F'ye) her birinde temel etkileri ve etkileşim etkilerini belirlemektir. Bu kutudaki her tablo ya da grafikte, her bağımsız değişken etkisinin, diğer bağımsız değişken düzeyine bağlı olarak değişip değişmediğini belirleyin. Yani, etkileşim var mıdır?

Daha sonra, diğer bağımsız değişken(ler)i göz önüne almadan (yani, onu yok sayarak), her bir değişkenin bir etki yapıp yapmadığını kontrol edin. Yani, bağımsız değişkenlerin birisinin veya ikisinin bir temel etkisi var mıdır? Tablodaki verileri (Şekil 8.2) grafiğe ve grafikteki verileri (Şekil 8.3 ve 8.4) tablolara dönüştürmeye gayret ederseniz, bu alıştırmadan en üst düzeyde yararlanırsınız. Bu alıştırmanın amacı, karmaşık bir desenin sonuçlarını betimlemenin çeşitli yollarını iyi bir düzeyde kullanmanızı sağlamaktır.

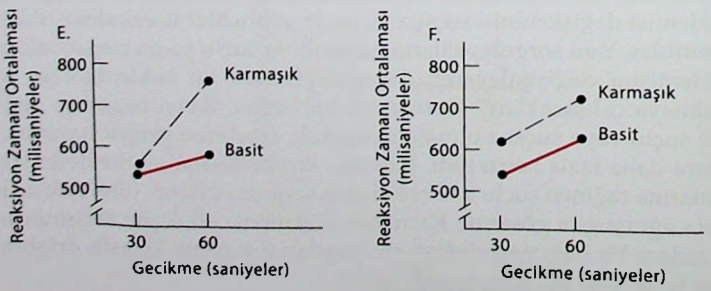
ŞEKİL 8.2 Görev zorluğu ve anksiyete düzeyinin bir işlevi olarak doğru tepkilerin ortalama sayısı

A.	Görev Zorluğu		B.	Görev Zorluğu	
Anksiyete Düzeyi	Kolay	Zor	Anksiyete Düzeyi	Kolay	Zor
Düşük	3.3	3.3	Düşük	6.2	3.1
Yüksek	5.6	1.2	Yüksek	4.2	1.1

ŞEKİL 8.3 Medya türü ve içeriğinin bir işlevi olarak agresif tepkilerin ortalama sayısı



ŞEKİL 8.4 Örüntü karmaşıklığı ve gecikmenin bir işlevi olarak ortalama reaksiyon zamanı



Karmaşık bir desende betimleyici istatistiklerden bir özet rapor hazırlamanın üç yaygın yolu vardır: Tablolar, bar grafikler ve çizgi grafikler. Bu tür tablo ve şekillerin hangi şekilde sunulacağına karar verme ölçütleri Bölüm 13'te açıklanmıştır. Genel olarak, tablolar herhangi bir karmaşık desende kullanılabilir ve deneydeki kesin değerlerin bilinmesi gerektiği zaman her koşul için en faydalı yöntemlerden birisidir. Öte yandan, çubuk grafikler ve çizgi grafikler, kesin değerleri belirtmeden, sonuçları özet hâlinde göstermek için çok yararlıdır. Çizgi grafikler özellikle karmaşık desenlerin sonuçlarını betimlemede yararlıdır, çünkü bir etkileşim çizgi grafikte kolaylıkla görülür. *Grafikte paralel olmayan çizgiler, bir etkileşimin olduğunu gösterir; paralel çizgiler ise etkileşimin olmadığını gösterir.* Örneğin, Şekil 8.1'e bakınız.

Bir 2 x 2 deseninin sonuçları bir tabloda özetlendiği zaman bir etkileşimin olup olmadığını değerlendirmenin en kolay yolu, *çıkarma yöntemi*dir. Çıkarma yöntemi, tablonun her sırasındaki (veya her sütunundaki) ortalamaların farklarını karşılaştırmak demektir. Eğer fark varsa, bir etkileşimin olması olasıdır. Çıkarma yöntemi uygulanırken, farkların aynı yönde hesaplanması zorunludur. Örneğin, Tablo 8.3'teki verilerde çıkarma yöntemini kullanmak için, suçlu beklentisi koşulundaki (5.64 – 7.17 = –1.53) şüpheli statüsünün iki düzeyine (suçluya ve suçsuza) ait ortalamaları çıkarabilirsiniz ve sonra, aynı işlemi suçsuzluk beklentisi koşuluna ilişkin olarak da yapabilirsiniz (5.56 – 5.85 = –0.29). Bulunan farkın işareti dikkatli bir şekilde kaydedilmelidir. Çıkarma yöntemi bu farklar arasında fark olduğunu ve böylece iki değişken arasında bir etkileşimin olası olduğunu gösterir. Çıkarma yöntemi, sadece bağımsız değişkenlerden birinin iki düzeyi varsa kullanılabilir. Karmaşık desenlerde, bağımsız değişkenlerin ikisinin de üç ya da daha fazla düzeyi olduğu zaman, etkileşim etkilerini belirtmek için grafikler kullanılmalıdır.

Üç Bağımsız Değişkenli Karmaşık Desenler

Karmaşık desenlerin gücü ve karmaşıklığı, deneydeki bağımsız değişkenlerin sayısı ikinin üzerine çıktıkça dereceli olarak artar. İki değişkenli (faktörlü) bir desende, sadece bir etkileşim etkisi gözlenir, oysa üç değişkenli bir desende her bir değişken iki diğer bağımsız değişkenin her birisi ile etkileşebilir ve üç bağımsız değişkenin üçü birlikte etkileşebilir. Böylece, iki değişkenli bir deney üç değişkenli bir deneye dönüştürüldüğü zaman, dört farklı etkileşim etkisi elde etme olasılığı ortaya çıkar. Üç bağımsız değişken A, B ve C olsun; bu durumda üç değişkenli desen A, B ve C temel etkilerini, iki yönlü etkileşim etkilerini (A x B, A x C, B x C) ve üç yönlü etkileşim etkisini (A x B x C) test etme olanağı sağlar. Üç bağımsız değişkeni olan bir deneyin yeterliliği açıktır. İşyerindeki aşırı kilolu kişilere karşı ayrımcılığı araştıran bir deney, karmaşık desenlerin ne kadar güçlü desenler olabileceğine dair açık bir fikir verecektir.

Bir 2 x 2 x 2 karmaşık desen deneyindeki katılımcılar kurgusal bir iş görüşmesi videosu izledi (Pingitore, Dugoni, Tindale & Spring, 1994). Bu durumda her birinin iki düzeyi olan üç bağımsız değişken, sekiz deneysel koşul ortaya çıkardı. İş görüşmesini izledikten sonra, katılımcılar, başvuran kişiyi işe alıp

almayacaklarını 7 dereceli bir ölçek kullanarak değerlendirdiler (1= *kesinlikle işe alınamaz* ve 7 = *kesinlikle işe alınır*.). Bu değerlendirme bağımlı değişkendi.

İlk Bağımsız Değişken: İşe Başvuranın Kilosu Araştırmacılar aşırı kilolu kişilere ilişkin damgalama (stigma) ile ilgilendiler; bu nedenle videoda görülen başvuranın kilosunu, normal ve aşırı kilolu olmak üzere iki düzeyde değişimlediler. İşe başvuran kişi rolünde normal kilolu profesyonel aktörler oynadı. Bir senaryoya uygun şekilde yaptıkları iş görüşmesi, normal kilolarıyla videoya alındı. Aşırı kilolu koşulda, aktörler aynı senaryoyu tekrar oynadılar ama bu defa makyaj yaparak ve plastik protez kullanarak %20 daha kilolu göründüler. Videodaki aktörlerin ve senaryonun aynı kaldığına dikkat edin. Katılımcılar, normal ağırlıktaki veya aşırı kilolu başvuranın olduğu koşullara seçkisiz atandılar (seçkisiz gruplar deseni).

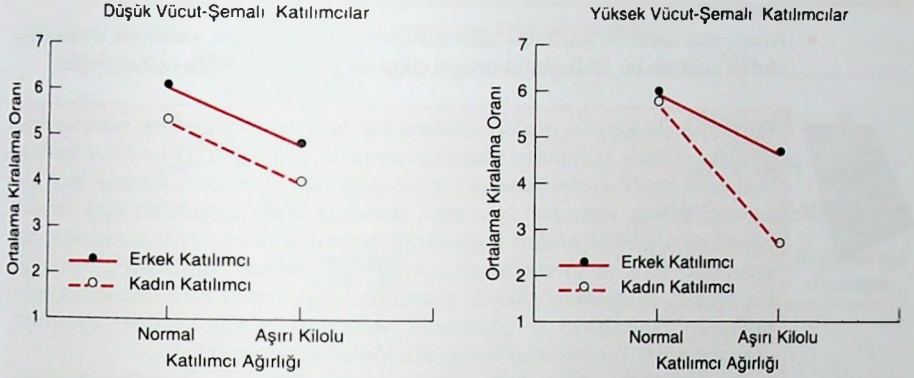
İkinci Bağımsız Değişken: Başvuranın Cinsiyeti Araştırmacılar aynı zamanda aşırı kilolu kişilere karşı geliştirilen damgalamanın, cinsiyete göre farklılaşıp farklılaşmadığını da test ettiler. Böylece, katılımcılar bir kadın veya erkek başvuranın bulunduğu video filmi koşuluna seçkisiz atandılar. Bu iki bağımsız değişkenin faktöryel birleşimi ile, *başvuranın cinsiyeti* ve *başvuranın ağırlığı* olmak üzere, deneyde videoya alınmış iş görüşmelerinin gösterildiği dört farklı seçkisiz grup oluştu: Normal kadın, kilolu kadın, normal erkek, kilolu erkek. Katılımcılar dört görüşmeden birisini izlemek üzere seçkisiz olarak atandı.

Üçüncü Bağımsız Değişken: Vücut Şeması Katılımcılar deneyden önce bir anket doldurdular. Bu ankette, kendi vücutları hakkındaki değerlendirmeleri ve bu değerlendirmenin, onların benlik kavramı içindeki önemi ölçüldü. Bu ölçümlere dayanan puanlarla bir *vücut-şeması* değişkeni oluşturuldu. Bu doğal gruplar deseni kullanılarak, katılımcılar düşük veya yüksek olmak üzere iki gruba ayrıldılar. Yüksek vücut–şeması olan katılımcılar, kendi vücutları ile vücutlarının kendi benlik kavramı içindeki önemi açısından daha fazla endişe gösteren katılımcılardı.

Üç Faktörlü Bir Desenin Sonuçlarını Anlamak Pingitore ve arkadaşlarının deneyinde bu üç değişkene ait sonuçlar Şekil 8.5'tedir. Üç bağımsız değişkenli bir deneyde ortalamaları göstermek birden fazla paneli olan bir grafik gerektirir. Şeklin bir paneli, iki değişkenin, üçüncü değişkenin bir düzeyindeki sonuçlarını gösteriyor ve diğer panel, aynı iki değişkenin, üçüncü değişkenin ikinci düzeyindeki sonuçlarını gösteriyor.

Şimdi temel etki ve basit (iki yönlü) etkileşime aşina olduğunuza göre, bir üç faktörlü ya da üç yönlü etkileşimi anlamaya yoğunlaşabiliriz. Şekil 8.5 gösteriyor ki bir katılımcının kilosuna ve cinsiyetine ait iki yönlü etkileşim, sadece kendi vücutları hakkında çok endişeli olan katılımcılarda ortaya çıkıyor. Yani, vücut-şema değişkeni yüksek olanlar (sağdaki panel) aşırı kilolu kadın başvuranlara düşük puanlar verdiler ama normal kilodaki kadın ve erkek

ŞEKİL 8.5 Bir 2x2x3 karmaşık desende etkileşim etkisinin gösterilmesi



başvuranlara aynı puanları verdiler. Öte yandan, vücut-şeması değişkeni düşük olan katılımcılar (soldaki panel), aşırı kilolu katılımcılara daha düşük puanlar verdiler ama kadın ve erkek katılımcılara ilişkin değerlendirmeler arasındaki fark, katılımcı ağırlık değişkeninin her iki düzeyi için de aynıydı.

Şekil 8.5'te gösterilen Pingitore ve arkadaşlarının (1994) bulgularını özetlemenin bir yolu, katılımcıların cinsiyet ve ağırlık bağımsız değişkenlerinin etkileşiminin, katılımcıların beden şemasına dayandığını söylemektir. Bu tür bir bulguyu üç yönlü (ya da üçlü) etkileşim diye adlandırırız. Üç yönlü bir etkileşim ortaya çıkarsa, sonuçları betimlerken üç bağımsız değişkenin üçü de dikkate alınmalıdır. Genelde, iki bağımsız değişken varsa, bir bağımsız değişkenin etkisi, ikinci bağımsız değişkenin düzeyine bağlı olarak farklılaşıyorsa bir etkileşim ortaya çıkar. *Bir karmaşık desende üç bağımsız değişken olduğu zaman, üç yönlü bir etkileşim, iki bağımsız değişkenin etkileşimi, üçüncü bağımsız değişken düzeyine bağlı olarak değiştiğinde ortaya çıkar.* Şekil 8.5'teki sonuçlar bunu gayet iyi göstermektedir. İlk iki bağımsız değişkene ait sonuçlar örüntüsü (başvuranın vücut ağırlığı ve cinsiyeti), üçüncü değişken düzeyine (katılımcının vücut şeması) bağlı olarak değişiyor. Pingitore ve arkadaşları üçüncü bir bağımsız değişken düzeyi olarak vücut şemasını da çalışmaya katarak, katılımcının ağırlığına dayalı, sadece cinsiyet ve ağırlık bağımsız değişkenlerinin kullanılacağı duruma kıyasla çok daha iyi ayırt eden bir yaklaşım sağladılar.

KARMAŞIK DESENLERİN ANALİZİ

- İki bağımsız değişkenli bir karmaşık desende, çıkarımsal istatistikler üç etkiyi test etmek için kullanılır: Her bağımsız değişkene ait temel etkiler ve iki bağımsız değişken arasındaki etkileşim.

- Betimleyici istatistikler, çıkarımsal istatistiğin sonuçlarını yorumlamak için gereklidir.
- Araştırmacıların bir karmaşık deseni nasıl yorumladıkları, verilerde istatistiksel olarak anlamlı bir etkileşimin ortaya çıkıp çıkmadığına bağlı olarak değişir.



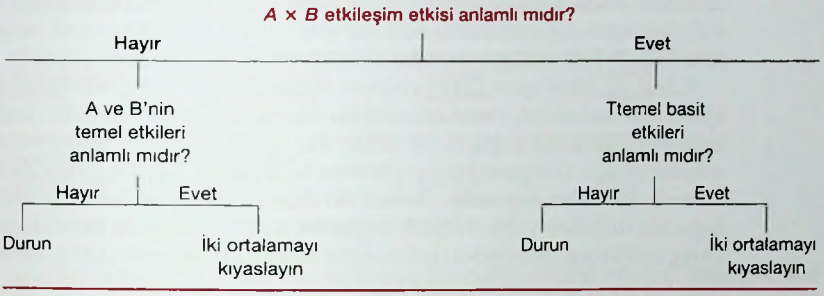
Karmaşık desenlerin analizi, sadece bir bağımsız değişkeni olan deneylerin analizinde kullanılan mantıkla aynıdır (Bölüm 6, 11 ve 12'ye bakın). Veri analizinde veriyi hatalar ve uç değerler açısından kontrol ettikten sonraki aşama, sonuçları ortalama, standart sapma, standart hata ve etki büyüklüğü gibi ölçümleri kullanarak betimlemektir. Daha sonra sıfır hipotezinin testi ve güven aralıkları gibi çıkarımsal istatistikler, herhangi bir etkinin istatistiksel olarak güvenilir olup olmadığını belirlemek için kullanılır. Betimleyici ve çıkarımsal istatistiklerin sonuçlarına dayanarak araştırmacılar, bu sonuçlar hakkında iddiada bulunabilirler.

Bu bölümün geriye kalan kısmındaki göreviniz, bir araştırmacının etkileşim etkisini ve temel etkiyi yorumladığı şekilde, karmaşık desenlere uygulanan veri analizini anlamaktır. Bu bölümde verilecek olan "Karmaşık Desenlerin Analizi" başlıklı girişi okumanız ve daha sonra Bölüm 12'de bu konunun tartışmasını tekrar gözden geçirmeniz sizin için faydalı olabilir. Bu bölümlerin her ikisinde de, meselenin zor yanından ziyade, bu analizlerin mantıkî gerekçesi ön planda tutulmuştur. Neyse ki, bilgisayarlar karmaşık desenlerde üretilen verilerle ilgili detaylı hesaplamaları yaparak zaman kazandırmaktadır. Ama bilgisayarlar bu hesaplamaların sonuçlarını yorumlayamaz. İşte siz burada devreye girersiniz. Yavaş ilerleyin; materyali dikkatli inceleyin ve metindeki betimlemeye eşlik eden tablo ve şekilleri incelediğinizden emin olun.

İki değişkeni olan bir karmaşık desenin üç potansiyel *sistemli değişkenlik kaynağı* olduğunu açıkladık; iki temel etki ve bir etkileşim. Karmaşık desenler sıfır hipotezi (F testi) ve güven aralıkları kullanılarak analiz edilir (Bölüm 12'ye bakınız). Bir karmaşık desende istatistiksel olarak anlamlı etki demek (herhangi bir analizde olduğu gibi), sıfır hipotezi çerçevesinde .05'in altında kabul edilen bir olasılık demektir (Bölüm 6'ya bakınız). Çıkarımsal istatistik testleri, bir etkileşimin gerçekte meydana gelip gelmediğini belirlemek amacıyla, betimsel istatistiklerle birlikte kullanılır. Verileri etkileşim açısından inceledikten sonra araştırmacılar, her bir bağımsız değişken için temel etkinin varlığını araştırabilirler.

Sadece bir bağımsız değişkeni olan bir deneyde olduğu gibi, bir karmaşık desende de sonuçları yorumlamak için ek analizler gerekebilir. Örneğin, bir araştırmacı ortalamalar arasındaki farklılıkları test etmek için güven aralıklarını kullanabilir. Bölüm 12'de böyle bir yaklaşımı örnekle gösterdik. Karmaşık desenlere ilişkin analiz planı, deneyde istatistiksel olarak anlamlı bir etkileşim olup olmadığına bağlı olarak değişir. Tablo 8.4 bir karmaşık deseni yorumlarken, etkileşim etkisinin ortaya çıktığı ve çıkmadığı durumlarda kullanılabilecek temel ilkeleri vermektedir. Bir deney betimleyerek, bu tabloda her iki yolu da göstereceğiz; bu deneyde istatistiksel olarak anlamlı bir etkileşim var ve daha sonra betimleyeceğimiz araştırmada ise istatistiksel olarak anlamlı bir etkileşim yoktur.

TABLO 8.4 İKİ FAKTÖRLÜ DESENLERİN ANALİZİ İÇİN YÖNLENDİRİCİ ÖNERİLER



Etkileşimin Olduğu Bir Durumda Analiz Planı

- Bir karmaşık desenin analizinde anlamlı bir etkileşim ortaya çıkarsa, etkileşimin kaynağı, basit temel etkilerin analizi ve iki ortalamanın karşılaştırılması yapılarak belirlenir.
- Basit temel etki, bir bağımsız değişkenin, ikinci bir bağımsız değişkenin bir düzeyindeki etkisidir.

Karmaşık bir desende etkileşim analizini anlayabilmek için, damgalanmış bireyler hakkındaki ön yargıyı anlamaya yönelik çağdaş bir yaklaşımı inceleyeceğiz. Sosyal psikologlar ön yargının bir etkisinin, damgalanmış gruplardaki insanların (örneğin; etnik azınlıklar, gay ve lezbiyenler, obez kişiler) toplum içinde değersiz görüldüklerine dair inanç sistemleri geliştirmek olduğunu öne sürerler. Damgalanmış kişiler, bu "sosyal kimlik tehdidi" ile, kendi çevrelerinde onlara olumsuz bakıldığını gösteren ipuçlarına karşı özellikle tetikte olmalarına yol açan beklentiler geliştirirler (Kaiser, Vick ve Major, 2006). İpuçlarına yönelik bu dikkat, *bilinçli* bir şekilde ortaya çıkabilir; bu durumda bireyler, damgalama ipuçlarına karşı özel bir dikkat gösterdiklerinin farkındadırlar. Bununla birlikte, daha yakınlarda araştırmacılar, sosyal kimlik tehdidinin insanların, farkında olmaksızın, olası damgalama ipuçlarına ne derece duyarlı olduklarını test ettiler.

Bilinç dışı dikkati incelemenin bir yolu "Duygusal Stroop görevi"dir. Bu görevin orijinal şekli size tanıdık gelebilir; katılımcılardan, kelimelerin hangi renkte yazılmış olduklarını söylemeleri istenir. Stroop görevi okumanın otomatik olduğunu (en azından yetişkinlerde) göstermek için tasarlanmıştır. İnsanların, renklerin adını söylemeye çalışırken, kelimelerin anlamını ihmal etmeleri çok zordur. Bu otomatik sürecin etkisinin en dramatik hâli, farklı renkte yazılmış renk isimlerinin olduğu koşulda gösterilmiştir (örneğin, "kırmızı" kelimesinin mavi ile yazılması). Böyle bir uyumsuzluk koşulunda katılımcıların rengin adını söylemeleri daha uzun zaman alır, çünkü kelimenin anlamı ile rengi karışır. Daha ileri çalışmalar gösteriyor ki kelimeler, katılımcının sunulan kelimenin farkında olmasının olanaksız olduğu kadar hızlı (örneğin, 15 milisaniye) sunulduğu zaman bile bu etki ortaya çıkıyor.

Duygusal Stroop görevinde, renk kelimelerinin yerine, katılımcıların endişeleriyle yakından ilişkili kelimeler kullanıldı. Örneğin, fobisi olan insanlarda bilinçli

olmayan dikkati inceleyen bir deneyde "yılan" ve " örümcek" gibi kelimeler kullanılabilir. Fobik katılımcıların bu kelimelerin renklerini belirleme süresi, nötr kelimelerinkinden daha uzundur; bu kelimeler farkında olunamayacak kadar kısa bir süre (bilinçli farkındalığın dışında) sunulduğu zaman bile sonuç aynıdır.

Kaiser ve arkadaşları (2006) cinsiyet ayrımcılığı ile damgalandıkları beklentisi içinde olan kadınların, cinsel ayrımcılıkla ilişkisiz kelimelere kıyasla, cinsel ayrımcı kelimelere karşı daha güçlü bir bilinç dışı dikkat gösterip göstermeyeceklerini araştırmak için Duygusal Stroop görevini kullandılar. Bir 2 x 3 faktörlü karmaşık desende 35 kadını test ettiler. Birinin iki düzeyi, diğerinin ise üç düzeyi olan iki bağımsız değişken vardı. Bağımlı değişken, tepki zamanıydı; bu da katılımcıların Duygusal Stroop görevindeki kelimelerin rengini belirlemeleri için geçen zamandır.

İlk Bağımsız Değişken: Sosyal Kimlik Cinsel ayrımcılık tehdidinin, bilinç dışı dikkati etkileyip etkilemediğini belirlemek için araştırmacılar, bir seçkisiz gruplar deseninde sosyal kimliği iki düzeyde değişimlediler: Tehdit ve güvenlik. Katılımcılara bir bilgisayar görevini tamamladıktan sonra, bir grup projesinde (gerçekte olmayan) bir erkekle birlikte çalışacakları söylendi ve bu erkeğin kişilik özellikleri hakkında bilgi verildi. *Tehdit* koşulunda, erkek cinsiyet ayrımı yapan birisiydi (örneğin, "Ben kadın bir patronla kesinlikle çalışamam çünkü kadınlar çok duygusaldır."). *Güvenli* koşulda ise erkek cinsiyet ayrımı yapmayan birisi olarak temsil edildi.

İkinci Bağımsız Değişken: Kelime Türü 2 x 3 faktörlü desende ikinci bağımsız değişken *kelime türü* idi ve üç düzeyi vardı: Sosyal kimlik tehdidi, hastalık tehdidi, ve tehdit yok. Bu değişken, tekrarlı ölçümler deseninde değişimlendi; böylece tüm katılımcılar tam bir karşıt dengelemenin yapıldığı bir sırada üç kelime türünün hepsini aldılar. *Sosyal kimlik tehdidi* olan kelimeler içerik olarak cinsel ayrımcılık gösteren kelimelerdi (örneğin, *fahişe*, *meme* gibi). *Hastalık açısından tehditkâr* kelimeler (örneğin, *kanser* gibi), tehdit koşulunda yer alan kadınların, sadece sosyal kimliği tehdit eden kelimelere değil, genel olarak tehdit içeren kelimelere karşı daha fazla dikkat edip etmediklerini anlamak için verildi. Aynı zamanda kontrol koşulu olan *Tehdit yok* koşulundaki kelimeler, evde yaygın olarak bulunan nesnelerdi (örneğin, *süpürge*, *perde* gibi). Kelimeler farklı renklerde (kırmızı, sarı, mavi, yeşil) ve eşik altı hızda (15 msn) sunuldu. Katılımcıların görevi, kelimelerin yazıldığı rengin adını söylemekti. Testler katılımcıların, hangi kelimelerin sunulduğunun farkında olmadıklarını gösterdi.

2 x 3 Etkileşim Etkisini Anlamak Kaiser ve arkadaşlarının tahmin ettiği gibi, iki bağımsız değişken arasında etkileşim ortaya çıktı. Hastalık açısından tehditkâr kelimelere ve nötr kelimelere kıyasla, sosyal kimlik açısından tehditkâr olan kelimeler sunulduğunda, tehditkâr koşulda yer alan kadınların (Tablo 8.5'teki ilk sıra) renkleri adlandırması daha uzun zaman aldı. Renkleri adlandırmak için daha uzun tepki zamanları, kadınların kelimelere daha uzun eşik altı dikkat yönelttiğini gösteriyor. Bu nedenle cinsel ayrımcı bir erkek çalışma arkadaşıyla temasa geçme beklentisi olan kadınlar, sosyal kimliklerini tehdit eden kelimelere daha fazla eşik altı dikkat harcamışlardır. Tersine, güvenli koşuldaki (Tablo 8.5'in ikinci sırası) cinsel ayrımcı olmayan bir erkekle çalışma beklentisi içinde

TABLO 8.5 SOSYAL KİMLİK VE KELİME TÜRÜNÜN (EŞİKALTI SUNUM) BİR İŞLEVİ OLARAK ORTALAMA TEPKİ ZAMANLARI (MS)

	Kelime Türü		
Sosyal Kimlik Koşulu	Sosyal Kimlik Açısından Tehditkâr	Hastalık Açısından Tehditkâr	Tehditkâr Olmayan
Tehditkâr Olan	598.9	577.7	583.9
Güvenli olan	603.0	615.0	614.5

Kaiser ve ark. (2006) verisinden adapte edilmiş olan veri

olan kadınların üç farklı tür kelimeye harcadığı dikkat süreleri arasında fark çıkmadı. Bir etkileşim etkisi var, çünkü kelime türü değişkeninin etkisi, sosyal kimlik değişkeninin düzeylerine (tehditkâr, güvenli) bağlı olarak değişmektedir. Sıfır hipotezi anlamlılık testi kullanan bu sonuçların çıkarımsal istatistikleri, etkileşim etkisinin istatistiksel olarak anlamlı olduğunu doğruladı.

Bir veride bir etkileşim etkisi bir kez doğrulandıktan sonra, etkileşimin spesifik kaynağı ek istatistiksel teknikler kullanılarak belirlenir. Tablo 8.4'te belirtildiği gibi, anlamlı bir etkileşimin kaynağını belirlemede kullanılan spesifik testler, basit temel etki ve iki ortalamanın karşılaştırılması olarak adlandırılır (Bölüm 12'ye bakınız).

Temel Kavram

Basit bir temel etki bir bağımsız değişkenin, ikinci bir bağımsız değişkenin *bir düzeyindeki* etkisidir. Sosyal-kimlik deneyinin sonuçlarına dönerek, basit temel etkilerin kullanımını örnekle gösterebiliriz. Tablo 8.5'te beş basit temel etki vardır; sosyal kimliğin iki düzeyinin her birindeki kelime türü etkisi ve kelime türünün üç düzeyinin her birindeki sosyal kimlik etkisi. Araştırmacılar eşik altı dikkat etkisinin (üç farklı kelime türü ortalaması arasındaki fark) tehdit koşulunda ortaya çıkacağını, güven koşulunda ortaya çıkmayacağını yordadılar. Bu yüzden, sosyal kimlik bağımsız değişkeninin her düzeyindeki kelime türü basit temel etkilerini test etmeyi seçtiler. Beklendiği gibi, kelime türünün basit temel etkisinin, tehdit koşulunda (üç ortalama arasındaki fark; 598.9, 577.7 ve 583.9) anlamlı olduğunu buldular. Yine de, kelime türünün basit temel etkisi, güven koşulunda anlamlı değildi (üç ortalama arasında fark yoktur; 603.0, 615.0 ve 614.5).

Kaiser ve arkadaşlarının deneyindeki kelime tipi bağımsız değişkeninde olduğu gibi, basit bir temel etkide üç veya daha fazla ortalama test edildiği zaman, basit temel etkinin kaynağını belirlemek için bir defada iki tanesi test edilen ortalamaların karşılaştırılması yapılabilir (Bölüm 12'ye bakınız). İlk olarak, güvenli koşul için ek analiz gerekmez, çünkü kelime türünün basit temel etkisi istatistiksel olarak anlamlı değildir. Bir sonraki aşama, basit temel etkinin anlamlı olduğu yerde, yani tehdit koşulundaki ortalamaları dikkatle analiz etmektir.

Kaiser ve arkadaşları, bir defada iki ortalamayı ele aldıkları analizlerinde, tehdit koşulundaki kadınlar için hem beklenen hem de beklenmeyen bir etki bildirdiler. Beklendiği gibi, ortalama tepki zamanları sosyal-kimliği tehdit eden kelimelerde, hastalık açısından tehditkâr olan kelimelere kıyasla, daha uzundu. Ama beklenmedik şekilde, nötr kelimeler sosyal-kimliği tehdit eden kelimelerle veya hastalık tehdidi içeren kelimelerle karşılaştırıldıkları zaman,

ortalama tepki zamanları arasında fark çıkmadı. Bu önemli bir soruyu gündeme getirmektedir: Kadınlar neden sosyal kimliği tehdit eden kelimelere verdikleri eşik altı dikkati, tehdit içermeyen kelimelere de vermektedirler? Kaiser ve arkadaşları, kadınların cinsel ayrımcılık gösteren bir erkekle çalışma beklentisinde oldukları zaman, tehdit içermeyen koşulda yer alan ev eşyalarına ait kelimeleri (örneğin, *soba, süpürge, mikrodalga fırın*) farkında olmadan temizlik ve yemek pişirme gibi çoğunlukla kadınlara özgü ev işleri ile ilişkilendirmiş olabilecekleri sonucuna vardılar. Kaiser ve arkadaşlarına (2006) göre, "Tehdit içermeyen kelimeler, en iyi karşılaştırmayı sağlamamış olabilirler." (s. 336). Onların bu beklenmeyen bulguya ilişkin yorumu, bir deneyin yorumlanması açısından, deneyin nasıl yapıldığının ve verinin nasıl analiz edildiğinin ne kadar önemli olduğunu gösterir.

Bir etkileşim etkisi bir defa analiz edildikten sonra, araştırmacılar aynı zamanda her bağımsız değişkenin temel etkilerini incelerler. Yine de, bir etkileşim ortaya çıktığını biliyorsak, temel etkiler daha az ilgimizi çeker. Örneğin, bu deneydeki etkileşim, bize farklı kelime türlerine yöneltilen eşik altı dikkatin, sosyal kimlik tehdidinin düzeyine bağlı olarak farklılaştığını söylemektedir. Biz bunu öğrendikten sonra, güvenli koşuldaki kadınların tehdit koşulundaki kadınlara kıyasla, tüm kelime türlerinde daha uzun tepki zamanı gösterip göstermediklerini öğrenerek fazla bir bilgi kazanmış olmayacağız. Kaiser ve arkadaşlarının çalışmalarında, kelime türü ve sosyal kimlik bağımsız değişkenlerinin temel etkileri istatistiksel olarak anlamlı değildi. Böyle olmakla birlikte, etkileşim etkisinin ve temel etkilerin ilgi çektiği deneyler vardır.

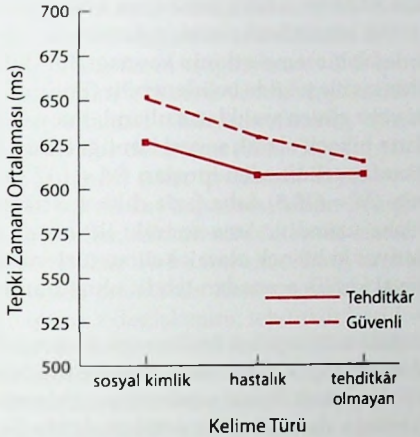
Etkileşimin Olmadığı Bir Durumda Analiz Planı

- Karmaşık bir desenin analizi bağımsız değişkenler arasındaki etkileşimin anlamlı olmadığını gösterirse, analizde bir sonraki basamak, değişkenin temel etkilerinin istatistiksel olarak anlamlı olup olmadığını belirlemektir.
- İstatistiksel olarak anlamlı bir temel etkinin kaynağı, iki ortalamayı karşılaştırarak ya da güven aralıkları kullanılarak hassas bir şekilde belirlenebilir.

Bir etkileşim etkisi istatistiksel olarak anlamlı *çıkmadığı* zaman, Kaiser ve arkadaşlarının (2006) yaptıkları sosyal kimlik deneyinin farklı bir kısmına ait sonuçları, bir karmaşık desen analizini incelemek için kullanabiliriz. Zaten betimlediğimiz sonuçlar *eşik altı* olarak sunulan kelimelere ilişkindi; öyle ki kelimelerin sunum hızı, katılımcıların kelimeleri algılayamayacakları kadar hızlıdır (15 msn). Bununla birlikte, bu deneydeki katılımcılar aynı zamanda *bilinçli* düzeyde sunulan kelimelerle de test edildi. Bilinçli-dikkat koşulunda, kadınlar kelimenin rengini söyleyerek tepki verene kadar ekrandaki kelimelere baktılar.[2]

[2]Zeki okuyucular Kaiser ve arkadaşlarının (2006) çalışmasının 2 (sosyal kimlik) x 3 (kelime türü) x 2 (kelime sunumu; eşikaltı, bilinçli) bir karmaşık (karma) desen olduğunu görebilir. Kelime sunumunun iki düzeyi, bir tekrarlı ölçümler deseni kullanılarak değişimlenmiştir. Bu bağımsız değişkenler arasındaki 2 x 3 x 2 etkileşimi istatistiksel olarak anlamlıdır. Kaiser ve arkadaşları etkileşimin kaynağını belirlemek için yaptıkları daha ileri analizde, 2 (sosyal kimlik) x 3 (kelime türü) etkileşimini ayrı olarak eşik altı ve bilinçli sunumda analiz ettiler. Burada açıklandığı gibi, 2 x 3 etkileşimi eşik altı sunumda istatistiksel olarak anlamlı, fakat bilinçli sunumda anlamlı değildi.

ŞEKİL 8.6 Etkileşim etkisi olmayan, fakat temel etkisi olan bir 2x3 karmaşık deseninin sonuçları (Dr. Cheryl R. Kaiser'den sağlanan veridir)



İki farklı kadın grubunda (tehditkâr koşul, güvenli koşul) üç kelime türü (sosyal kimliği tehdit eden, hastalık tehdidi yaratan, tehditkâr olmayan) için ortalama tepki zamanları Şekil 8.6'dadır. Etkileşimin yokluğu, şekilde görülebilir. Şekildeki iki çizgi mükemmel biçimde paralellik göstermiyor; ortalama tepki zamanlarının her iki grupta da yaklaşık aynı hızda azaldığı görülüyor. Çıkarımsal istatistik testleri etkileşim etkisinin anlamlı olmadığını doğruladı. Şekil 8.6'da gösterilen veriler, veri analizinin genel bir prensibini göstermektedir: *Betimleyici istatistikle gösterildiği gibi, bulgulara ait genel tablo deneyde bir etkileşimin olup olmadığına karar vermek için yeterli değildir. Etkilerin istatistiksel olarak güvenilir olup olmadığını doğrulamak için F–testi gibi çıkarımsal istatistik testleri yapılmalıdır.*

İstatistiksel olarak anlamlı bir etkileşim yoksa, bir sonraki aşama her bağımsız değişkenin temel etkisini incelemektir (Tablo 8.4' e bakınız). Kaiser ve arkadaşlarının bilinçli-farkındalık deneyindeki ortalamalar, temel etkilerin belirlenmesini kolaylaştırmak için Tablo 8.6'da tekrar verilmiştir. İki sosyal-kimlik koşulunun ortalamasını alarak, her bir kelime türüne ait ortalama tepki

TABLO 8.6 KİMLİK VE KELİME TÜRÜNÜN BİR İŞLEVİ OLARAK ORTALAMA TEPKİ ZAMANLARI

Sosyal Kimlik Koşulu	Kelime Türü			
	Sosyal Kimlik Açısından Tehditkâr Olan	Hastalık Açısından Tehditkâr Olan	Tehditkâr Olmayan	Sosyal Kimliğe İlişkin Ortalamalar
Tehditkâr koşul (n ∇ 18)	625.9	607.4	607.5	613.6
Güvenli koşul (n = = 16)	650.6	629.0	614.5	631.4
Kelime türüne ilişkin ortalamalar	637.5*	617.6*	610.8*	

Dr. Cheryl R. Kaiser tarafından sağlanan veri
*Ağırlıklı ortalamalar sosyal kimlik koşullarına ilişkin eşit olmayan örneklem boyutlarına dayanarak hesaplanmıştır

zamanlarını (yani, kelime türü değişkeninin temel etkisini) elde ettik. Bu ortalamalar sosyal kimliği tehdit eden kelimeler için 637.5, hastalık tehdidi içeren kelimeler için 617.6 ve tehdit içermeyen kelimeler için 610.8'dir. Kelime türünün temel etkisi istatistiksel olarak anlamlıydı. Üç veya daha fazla ortalamayı kapsayan anlamlı bir temel etkinin kaynağı, bir defada iki ortalama karşılaştırılarak daha hassas bir şekilde belirlenebilir (Bölüm 12'ye bakınız). Bu karşılaştırmalar *t*-testi veya güven aralıkları kullanılarak yapılabilir. Kaiser ve arkadaşları tüm kadınların bilinçli olarak sosyal kimliği tehdit eden ipuçlarına (M = 637.5), hastalık açısından tehdit eden ipuçları (M = 617.6) ve tehditkâr olmayan ipuçlarına kıyasla (M = 610.8) daha fazla dikkat ettiklerini gösterdi (örneğin, tepki zamanları daha uzundu). Ama sonraki iki koşul arasında fark yoktu. Bu sonuçlar gösteriyor ki bilinçli olarak kelime türlerinin farkında oldukları zaman, kadınlar sosyal kimlik açısından tehdit oluşturan kelimelere daha fazla dikkat ediyorlar.

Tablo 8.6'daki ortalamaları kullanarak da sosyal kimlik değişkeninin temel etkisini test edebiliriz. Kelime türü değişkeni üzerinden, tehdit (613.6) ve güvenlik (631.4) koşuluna ilişkin ortalamaları elde ederiz. Sosyal kimlik değişkeninin temel etkisi anlamlı değildi; ortalamalarda da görüldüğü gibi, kadınlarda tepki zamanları tehditkâr ve güvenli olan koşuldakine benzerdi. İki ortalamanın farklı gibi görünmesi, ortalamalar arasındaki farkların güvenilir olup olmadığını belirlemek için istatistiksel analize gerek olduğu düşüncesini desteklemektedir. Bir etkinin istatistiksel olarak anlamlı olup olmaması ortalamalara, örneklem boyutuna ve grup içi hata varyasyonuna bağlıdır.

Kaiser ve arkadaşlarının sosyal kimlik deneyinin analizi, istatistiksel olarak anlamlı bir etkileşim olmadığı zaman da karmaşık bir desenden öğrenilebilecek çok şey olduğunu gösteriyor.

ETKİLEŞİM ETKİLERİNİ YORUMLAMAK

Etkileşim Etkileri ve Teori Test Etme

- Teoriler sıklıkla iki ya da daha fazla bağımsız değişkenin etkileşiminin, davranışı etkilediği yordamasında bulunurlar; bu yüzden bu teorileri test etmek için, karmaşık desenlere ihtiyaç vardır.
- Bu teorilerin testleri bazen çelişik bulgulara yol açar. Etkileşim etkileri bu çelişik sonuçları çözümlemede faydalı olabilir.

Teoriler bilimsel yöntemde çok önemli bir rol oynar. Karmaşık desenler çoğunlukla araştırmacının teorileri test etme yeteneğini arttırır, çünkü her iki etkiyi de, hem temel etkiyi hem de etkileşim etkilerini test edebilir. Örneğin, Kaiser ve arkadaşları (2006) çevredeki ön yargı ipuçlarına dikkat edilmesine ilişkin hipotezleri test ettiler ki bu hipotezler sosyal kimlik teorisine dayanıyordu. Önceki araştırma, sosyal kimlikleri tehdit edildiğinde, bireylerin çevrelerindeki ön yargı ile ilişkili potansiyel ipuçlarının *bilinçli olarak* farkında olduklarını göstermiştir. Kaiser ve arkadaşları bu araştırmayı, "tehdit edilen bireylerin ön yargı içeren ipuçlarına farkında olmaksızın, *bilinçsizce* dikkat ettikleri" hipotezini test ederek genişletti. Karmaşık desen kullandıkları için Kaiser ve arkadaşlarının verileri, "güven" durumundaki kadınlarla kıyaslandığında, cinsiyetçi beklenti

içinde olan kadınların, cinsel ayrımcılık içeren kelimelere, diğer kelimelerden daha fazla bilinç dışı dikkat harcadıklarına dair kanıt sağlamaktadır. Onların verileri, ön yargı üzerine sosyal kimlik teorisini destekledi; bu teoriye göre, "damgalanmış grup üyeleri, değersiz olduklarına dair bir inanç sistemi geliştirirler ve bu beklentiler onların aşağılanmaya dair ipuçlarına karşı özellikle tetikte olmalarına neden olur". (Kaiser ve arkadaşları, 2006, s. 332).

Ayrıca Kaiser ve arkadaşları dikkat teorilerinin, dikkatin kapasitesinin sınırlı olduğunu belirttiklerine de dikkat çekti. Ön yargıyla karşılaşan insanların dikkati, sosyal kimliği tehdit eden ipuçlarına kayar ve bu yüzden diğer görevler için kullanabileceği dikkat kapasitesi azalır. Örneğin sınıf ortamında bir ön yargı algılayan öğrenciler, dikkatlerini farkında olarak ya da olmayarak bu potansiyel tehdide yöneltirler ve bu durum onların sınıf başarısını düşürebilir. Bununla birlikte, şu nokta önemlidir: Kaiser ve arkadaşları çalışmalarında sosyal kimlik tehdidi bağımsız değişkenini, tehditkâr koşul ve güvenli olan koşul olmak üzere iki düzeyde değişimledikleri için, bireylerin sosyal kimlik tehdidine karşı güvende olduklarına inandıkları koşulda, dikkat kaynaklarını potansiyel tehditlere yöneltmediklerini gösterebilmişlerdir. Bu bulgu, önyargıdan olabildiğince uzak çevreler oluşturmanın önemini bir defa daha gösterir.

Sosyal kimlik ve ön yargı gibi başlıkları içeren psikolojik teoriler genellikle karmaşıktır. Ön yargıyı incelemek için psikologlar örneğin, davranışsal, bilişsel ve duygusal süreçleri birey, grup ve toplum düzeyinde açıklamaya ihtiyaç duyarlar. Tahmin edebileceğiniz gibi, karmaşık teorileri test ederken kullanılan deneysel desenler çelişkili bulgulara yol açabilir. Örneğin, hipotetik bir ön yargı çalışması düşünün. Bu çalışma, düşük statülü bir grubun üyelerinin, sosyal-kimlik tehditlerine karşı bilinç dışı bir dikkat *göstermediğini* bulmuş olsun. Bu çelişkili görünen bulgu, damgalanmış bireylerin kimliklerine olan potansiyel tehditlere dikkat ettiğini öne süren bir önyargı teorisine nasıl katılır? Kaiser ve arkadaşlarının bulgularının gösterdiği gibi, bu bulgunun bir yorumu, sosyal kimlik koşulu değişkenini (güvenli veya tehditkâr) içerebilir. Bu çelişkili bulgu, hipotetik ön yargı araştırmasındaki katılımcıların, sosyal kimlik tehditlerine karşı kendilerini güvende hissettikleri ve bu yüzden küçümseyici potansiyel kaynaklara dikkatlerini harcamadıklarını öne sürerek de yorumlanabilirdi.

Çelişkili bulguları açıklamanın yaygın bir yolu şudur: Çelişkili bulguların potansiyel kaynaklarıyla ilişkili değişkenleri (örneğin; tehditkâr ve güvenli koşulları) araştırma desenine dâhil etmek. Daha genel bir ifadeyle teorileri test ederken karmaşık desenler, görünüşte çelişkili olan bulguların nedenlerini belirlemede çok yararlıdır.

Etkileşim ve Dış Geçerlik

- Bir karmaşık desende herhangi bir etkileşim gözlenmezse, her bir bağımsız değişkenin etkisi, diğer bağımsız değişkenin her düzeyine aynı şekilde genellenebilir; böylece bağımsız değişkenlerin dış geçerliği artar.
- Etkileşimin olması, bir bağımsız değişkenin etkisinin ortaya çıktığı koşulları açık hâle getirerek, bir bulgunun dış geçerliğinin sınırlarını belirler.

Bölüm 6'da, bir deneyde sadece bir bağımsız değişken olduğunda, bir araştırma bulgusunun dış geçerliğini sağlamanın yöntemlerini tartıştık. Dış geçerliği sağlamak için kısmî tekrar (partial replications) yapılabileceğini –yani, araştırma bulgularının nasıl genellenebileceğini– açıkladık. Ayrıca, alan deneylerinin araştırmacılara bağımsız değişkenleri gerçek ortamlarda nasıl inceleyeceklerine dair fırsat sağladığını tartıştık. Karmaşık desenlerde bir etkileşim etkisinin olması veya olmaması, bulguların dış geçerliğini belirlemede çok önemlidir.

Bir karmaşık desende etkileşim bulunmadığı zaman biliriz ki her bir bağımsız değişkenin etkileri, diğer bağımsız değişkenin düzeylerine genellenebilir. Örneğin, Kasinn ve arkadaşlarının (2003) sorgulayan kişilerin bir şüpheliyi sorgulamaları sırasındaki beklentilerine dair çalışmalarını düşünün. Şüphelinin aslında suçlu ya da suçsuz olmasından bağımsız olarak, suçsuz olduğu beklentisinde oldukları koşula kıyasla, suçlu olduğunu düşündükleri koşulda, sorgulayıcıların şüphelinin suçlu olduğu varsayımına dayanan daha fazla soru seçtikleri bulundu. Bu sonuç, "sorgulayıcının beklentisi" değişkeni ile "şüphelinin durumu" değişkenleri arasında herhangi bir etkileşim olmadığını gösterir. Bu nedenle, suç beklentisi durumunda sorgulayanın suç varsayımına dayanan sorular seçmesi, şüphelinin aslında suçlu ya da suçsuz olduğu durumlara genellenebilir.

Elbette bulgularımızı, deney koşullarının ya da sınırlarının dışına genelleyemeyiz. Örneğin, eğer kanun uygulayıcısı memurlar gibi diğer gruplar da test edilseydi, sorgulayıcının beklentileri ve şüphelinin statüsü arasında bir etkileşim olmaması, yine de suç varsayımına dayanan soruların seçiminin benzer olacağı sonucuna varmamıza izin vermeyecekti. Aynı şekilde, eğer sorgulayanların beklentilerine dair diğer değişimlemeler kullanılsaydı, aynı etkiler ortaya çıkar mıydı bilmiyoruz. Şunu da hatırlamalıyız ki istatistiksel olarak anlamlı bir etkileşim etkisi bulunmaması, etkileşim etkisinin gerçekten olmadığı anlamına gelmez; bunu gösterebilecek duyarlılıkta bir deney yürütememiş olabiliriz.

Etkileşim etkisinin olmaması her bir bağımsız değişkenin etkilerinin dış geçerliğini arttırmış olsa da, etkileşimin *olması* bir bulgunun dış geçerliğinin sınırlarını belirler. Örneğin, Kassin ve arkadaşları (2003 şüphelinin gerçekte suçsuz olduğu durumda bile,) şüphelinin suçlu olduğuna inanan sorgulayıcıların, kendisinden itiraf almak için daha fazla baskı yaptıklarını buldular. Bu etkileşim, sorgulayıcı beklentisinin, sorgulayıcının itiraf elde etmek için uyguladığı baskıya olan etkisinin dış geçerliğinin sınırlarını açık şekilde ortaya koyar. Bu bulguya dayanarak, sorgulayıcı beklentisinin itiraf almak için yapılan baskı üzerindeki genel etkisi için verilecek en iyi cevap "Duruma göre değişir." demektir; buradaki "durum" ise, şüphelinin gerçekten suçlu ya da suçsuz olmasıdır. Etkileşimin varlığı, dış geçerliğe ilişkin sınırları ortaya koyar ama aynı zamanda o sınırların neler olduğunu da belirler.

Bağımsız değişkenler arasında etkileşim olması olasılığı, bir bağımsız değişkenin davranış üzerinde bir etkisinin olmadığını söyleyeceksek temkinli olmamızı gerektirir. Davranışı etkileyen bağımsız değişkenlere, konuyla ilişkili

Temel Kavram

bağımsız değişkenler denilir. Genel olarak **ilişkili bağımsız değişken,** davranışı doğrudan etkiler (bir temel etki ile sonuçlanır) ya da ikinci bir bağımsız değişkenle birlikte çalışıldığında bir etkileşim doğurur. Davranışı etkileyen ve etkilemeyen etkenler arasında ayrım yapmak, davranışı açıklayan yeterli teoriler geliştirmek ve okul, hastane ve fabrika gibi uygulamalı alanlardaki sorunlarla ilgili etkin girişimler tasarlamak için gereklidir (Bölüm 9 ve 10'a bakınız)

Bir bağımsız değişkeni *ilişkisiz* olarak tanımlarken neden temkinli olmamız gerektiğinin birkaç nedeni vardır. Öncelikle, bir deneyde bir bağımsız değişkenin etkisi olmadığı gösterilmiş olsa da eğer bu bağımsız değişkenin farklı düzeyleri test edilmiş olsaydı, bu değişkenin etkisinin yine de gözlenmeyeceğini varsayamayız. İkincisi, eğer bir bağımsız değişkenin tek faktörlü bir deneyde hiç etkisi yoksa, bir karmaşık desen kullanıldığında, bir başka bağımsız değişkenle etkileşmeyeceği anlamına gelmez. Üçüncüsü, bir bağımsız değişkenin bir deneyde bir etkisi yoksa, farklı bağımlı değişkenlerle bir etkisinin ortaya çıkması mümkündür. Dördüncüsü, istatistiksel olarak anlamlı etkinin olmaması, etkinin olmadığı anlamına gelebilir de gelmeyebilir de. En azından biz, ilişkisiz bir değişkeni belirlemiş olduğumuza karar vermeden önce, deneyimizin duyarlılığı ve istatistiksel analizimizin gücü üzerinde düşünmek isteriz (İstatistiksel analizin gücüne ilişkin tartışma için Bölüm 12'ye bakınız). Şimdilik en iyisi, herhangi bir bağımsız değişkenin bir etkisinin olmadığı konusuna ön yargılı yaklaşmamanızı öneririz.

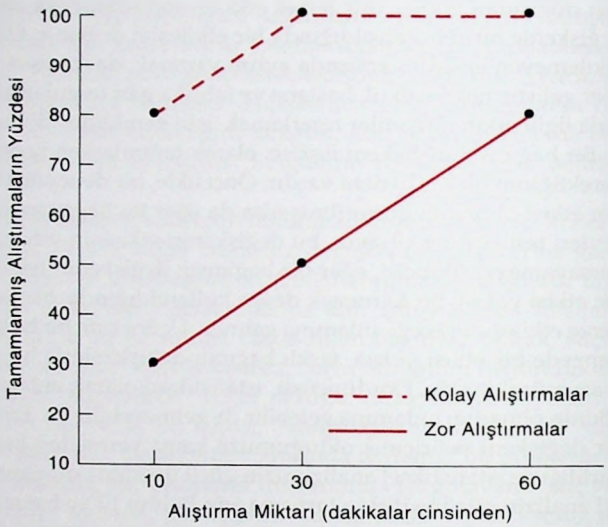
Etkileşim Etkileri ve Tavan ve Taban Etkileri

- Katılımcının performansı deneyin bir ya da daha fazla koşulunda en üst (tavan) ya da en alt (taban) düzeye ulaştığı zaman, etkileşim etkisinin sonuçları yorumlanamaz.

Bir fiziksel-fitness testi sırasında egzersizi arttırmanın performans üzerindeki etkilerini araştıran 3 × 2'lik bir deneyin sonuçlarını ele alın. Bu inandırıcı, ama hipotetik deneyde altı katılımcı grubu var. Katılımcılar öncelikle 10, 30 ya da 60 dakika kolay ya da zor egzersizler yaptılar. Daha sonra katılımcılara, yaptıkları bu egzersizleri içeren bir fitnes testi uygulandı. Bağımlı değişken her katılımcının 15 dakika içinde tamamlayabildiği egzersiz yüzdesi idi. Deney sonuçları Şekil 8.7'dedir.

Şekil 8.7'deki sonuç örüntüsü, klasik bir etkileşim etkisi gibi görünüyor. Egzersiz süresinin etkisi, kolay ve zor egzersizlerde farklı oldu. Egzersiz süresinin artması, zor egzersizlerde test performansını geliştirdi ama 30 dakikalık kolay egzersizden sonra test performansı aynı kaldı. Eğer standart analiz bu veriye uygulansaydı, çok muhtemeldir ki etkileşim etkisi istatistiksel olarak anlamlı olacaktı. Ne yazık ki bu etkileşim aslında yorumlanamaz. Kolay egzersiz verilen alıştırma grubunda, 30 dakikalık bir alıştırmadan sonra performans en üst düzeye ulaştı. Böylece 60 dakika alıştırma yapan grupta, bu noktanın ötesinde hiçbir gelişme görülmedi. Altmış dakika alıştırma yaptırılan katılımcılar egzersizden daha fazla faydalanmış olsalar bile, araştırmacı bu farkı seçtiği bağımlı değişkenle ölçemedi.

ŞEKİL 8.7 Tavan etkisinin gösterimi



Yukarıdaki deneyde tavan etkisi denilen bir genel ölçüm problemi gösterildi. Deneyin herhangi bir koşulunda performans ne zaman maksimuma ulaşsa, **tavan etkisi** (ceiling effect) tehlikesi vardır. Performans minimum olduğu zaman ise **taban etkisi** (floor effect) ortaya çıkar. Araştırmacılar tavan ve taban etkilerini engellemek için, koşullar arasında ölçülecek performans farklılıkları için geniş aralıklara izin veren bağımlı değişkenler seçerler. Örneğin, fitnes deneyinde katılımcılara, bir kimsenin herhangi bir testte tamamlayabileceği sayıdan çok daha fazla sayıda egzersiz verilmesi daha iyi olurdu. Bu durumda tavan etkisi tehlikesi olmaksızın, her koşulda tamamlanan egzersiz sayısı ortalaması, iki bağımsız değişkenin etkilerini değerlendirmek için kullanılabilirdi. Tavan etkilerinin karmaşık bir desen içermeyen deneylerde de aynı zamanda bir problem oluşturabildiğini göz önünde tutmak gerekir. Eğer fitnes deneyi sadece kolay egzersizleri içerseydi, deneyde hâlâ bir tavan etkisi olacaktı

Temel Kavramlar

Etkileşim ve Doğal Gruplar Deseni

- Doğal grupların neden farklılaştığına ilişkin bir teoriyi test ederken, araştırmacılar doğal gruplar değişkenlerine ilişkin nedensel çıkarımlar yapmak için karmaşık desenleri kullanırlar.
- Bir doğal gruplar değişkenini içeren nedensel bir çıkarım yapmada üç basamak şöyledir: Grup farklılıklarının neden ortaya çıktığını belirlemek, kuramsallaştırılan süreci gösteren bir bağımsız değişkeni değişimlemek değişimlenen değişken ve doğal gruplar değişkeni arasında bir etkileşim ortaya çıkıp çıkmadığını test etmektir.

Bölüm 6'da betimlenen doğal gruplar deseni, psikolojideki en popüler araştırma desenlerinden biridir. Bireysel farklar değişkenleri oluşturmak için cinsiyet, yaş, içedönüklük-dışadönüklük ya da saldırganlık gibi bazı özellikler açısından farklılaşan bireyler seçilerek gruplar oluşturulur. Araştırmacılar bu durumda bu bireysel farklar değişkenleri ile davranışın diğer yönleri arasında sistematik ilişkiler ararlar. Doğal gruplar deseni, bireylerin özellikleri ile performansları arasında korelasyonlar kurmada etkilidir. Bununla birlikte Bölüm 6'da da betimlediğimiz gibi, davranışın nedenleri hakkında sonuçlar çıkarma aşamasına gelindiğinde doğal gruplar deseni belki de en zorlayıcı desendir.

Doğal gruplar desenini yorumlamadaki zorluk, performanstaki farklılıklara, oluşturduğumuz gruplardaki insanların özelliklerinin neden olduğu sonucuna varmak istediğimizde ortaya çıkar. Örneğin, katılımcıların müzik eğitimlerine dayanılarak seçildikleri bir deneyi ele alın. Bir grup 10 ya da daha fazla yıl müzik eğitimi almış insanlardan, diğer grup ise özel bir müzik eğitimi almayan kişilerden oluştu. Her iki grup da 10 basit melodinin notalarını hatırlama yeteneği açısından test edildi. Sonuç şaşırtıcı değildi; eğitim almayanlara kıyasla müzik eğitimi alanlar çok daha iyiydi.

Bu sonuçlara dayanarak basit melodilere ilişkin belleğin, müzik eğitimi süresiyle orantılı şekilde değiştiği sonucuna varabiliriz. Fakat müzik eğitiminin daha iyi bir performansa *neden olduğu* sonucuna varamayız. Ama neden? On yıllık müzik eğitimi olanların, eğitimi olmayanlardan farklılaşabileceği olası birkaç neden vardır. Gruplar genel eğitimin miktarı ve türü, aile yapıları, sosyoekonomik durum ve dinlemiş oldukları müzikle ilgili deneyim miktarı ve türü açısından farklılaşabilir. Aynı zamanda, müzik eğitimi alanlar, genellikle bu tür bir eğitim almayanlara kıyasla daha iyi bir belleğe sahiptir ve onların performansı bu genel bellek yeteneğini yansıtabilir. Nihayet, müzik eğitimi almış olanların müziğe yetenekleri daha yüksek olduğu için performansları da daha yüksek çıkmış olabilir. Dolayısıyla, herhangi bir müzik eğitimi almamış olsalardı bile, müzik belleği görevinde daha başarılı olabilirlerdi. Kısaca, müzik eğitiminde gözlenen bellek performansındaki fark, müzik eğitimindeki bireysel farkların ötesinde birçok nedenden kaynaklanmış olabilir.

Doğal gruplar deseninde nedensel çıkarımlar yapma problemine potansiyel bir çözüm vardır (Underwood ve Shaughnessy, 1975). Bu çözümde anahtar, kritik bireysel farklar değişkeni hakkında bir teori geliştirmektir. Örneğin, Halpern ve Bower (1982) müzisyenler ve müzisyen olmayanlar arasında müzik belleğinin nasıl farklılaştığı ile ilgilendiler. Araştırmacılar "kümeleme" (chunking) adı verilen bir bellek kavramını kullanarak, müzik eğitiminin notaların işlenmesini nasıl etkilediğine dair bir teori geliştirdi. Şu 15 harflik diziyi ezberlemek istiyorsanız, kümeleme tekniği kullanarak bu görevi kolaylaştırabilirsiniz: HBOFBICNNUSAWWW. Bu dizileri ezberlemek zor görünüyor ama şu şekilde gruplarsanız, oldukça kolaylaştığını göreceksiniz: HBO-FBI-CNN-USA-WWW. Kümeleme tekniği, bu uzun harf dizisini, yine aynı harflerden oluşan beş anlamlı küçük diziye çevirerek, belleğin işini kolaylaştırdı.

Halpern ve Bower, müzik eğitiminin müzisyenlere, notaları, yukarıdaki kümelemede olduğu gibi, anlamlı "nota kümeleri" oluşturma olanağı sağladığı teorisini oluşturdu. Dahası, eğer kümeleme müzisyen olan ve olmayanların bellek performansları arasındaki farktan sorumlu ise, o zaman bu iki grup arasındaki fark, kötü müzikal yapısı olan melodilere kıyasla, iyi müzikal yapısı olan melodiler için daha büyük olmalıydı. Bu teoriyi test etmek için araştırmacılar müzikal yapı bağımsız değişkenini üç düzeyde değişimlediler. Nota yapıları görsel olarak birbirine benzeyen ama müzikal yapısı iyi, kötü ya da seçkisiz olacak şekilde düzenlenmiş basit melodi setleri hazırladılar.

Halpern ve Bower'in deneyinde, kritik test iki bağımsız değişken arasında bir etkileşim etkisi elde edip edemeyecekleri idi: Müzik eğitimi ve melodilerin türü. Yani, iyi bir müzikal yapı gösteren meleodilerde müzisyen olan ve olmayanlar arasındaki farkın en fazla olacağı, kötü yapı gösteren melodilerde daha az fark gözleneceği gözleneceği ve en az farkın ise seçkisiz melodilerde gözleneceği beklendi. Deneyin sonuçları, tam olarak böyle çıktı.

Elde edilen etkileşim etkisi, Halpern ve Bower'ın, müzisyen olanlarla olmayanlar arasındaki bellek performansı farkına ilişkin birkaç alternatif açıklamayı reddetmesini sağladı. Genel eğitimin miktarı ve türü, sosyoekonomik durum, aile arka planı ve iyi bir genel bellek yeteneği gibi özelliklerin, artık bu gruplar arasındaki performans farkını açıklama olasılığı yoktur. Bu potansiyel alternatif açıklamalar, iki grubun seçkisiz melodilerdeki performans farkının niçin farklılaşmadığını açıklayamaz. İşte etkileşim, bu tür basit korelasyonel açıklamaların olasılığını düşürür.

Doğal gruplar deseninde nedensel çıkarımlara varabilmek için gerekli işlemlerin, aşağıdaki aşamalarda yapılması gerekir.

Basamak 1: Bir Teori Geliştirin İlk basamak bireysel farklar değişkeninin grupların performansında niçin bir fark ortaya çıkarması gerektiğini açıklayan bir teori geliştirmektir. Örneğin, Halpern ve Bower melodileri bilişsel olarak organize etme (kümeleme) yöntemi açısından, müzisyen olanların ve olmayanların farklılaştığı kuramını öne sürdü.

Basamak 2: Değişimlenecek Bir İlişkili Değişken Belirlemek İkinci basamak, değişimlenebilen ve bu teorik sürecin ortaya çıkma olasılığını etkilediği varsayılan bir bağımsız değişken seçmektir. Halpern ve Bower, "müzikal yapının türü"nü, kümeleme kolaylığı ile ilişkili bir değişken olarak belirlemiştir.

Basamak 3: Etkileşimin Test Edilmesi Önerilen yaklaşımın en kritik yönü, değişimlenen değişken ile bireysel farklar değişkeni arasında bir etkileşim etkisi ortaya çıkarmaya çalışmasıdır. Böylece, değişimlenen ilişkili bağımsız değişken, her iki doğal grupta test edilir. Halpern ve Bower, 2 x 3'lü bir karmaşık

desende bireysel farklar değişkeni (müzisyenler ve müzisyen olmayanlar) ile değişimlenen değişken (müzikal yapının türü) arasında bir etkileşim buldular. Bu yaklaşım, üç bağımsız değişkenin etkileşimine ilişkin kestirimleri test ederek daha da güçlendirilebilir: Değişimlenen iki bağımsız değişken ve bireysel farklar değişkeni (bk. örneğin, Anderson & Revelle, 1982).

ÖZET

Karmaşık bir desen, iki ya da daha fazla bağımsız değişkenin etkisinin aynı deneyde araştırıldığı bir desendir. İki bağımsız değişken içeren karmaşık bir desen araştırmacıların her bir bağımsız değişkenin genel etkisini (temel etkisini) belirlemesine olanak tanır. Daha da önemlisi, karmaşık desenler iki bağımsız değişken arasındaki etkileşimi ortaya çıkarmada kullanılabilir. Etkileşim, bir bağımsız değişkenin etkisi, diğer bağımsız değişkenin düzeylerine bağlı olarak değiştiği zaman ortaya çıkar.

Olası en basit karmaşık desen, iki bağımsız değişkenin her birisinin ikişer düzeyinin araştırıldığı 2 x 2'li desendir. Faktöryel desende koşul sayısı, bağımsız değişkenlerin düzeylerinden ortaya çıkan koşul sayısına eşittir (örneğin, 2 x 3 = 6). 2 x 2'den daha fazla faktör içeren karmaşık desenler, davranışı anlamak için daha faydalı bile olabilir. 3 x 2, 3 x 3, 4 x 2, 4 x 3 gibi desenler elde etmek için bir ya da her iki bağımsız değişkenin ek düzeyleri eklenebilir. Aynı zamanda 2 x 2 x 2, 2 x 3 x 3 ve bunun gibi desenler elde etmek için ek bağımsız değişkenler eklenebilir. Üç bağımsız değişken içeren deneyler belirgin bir şekilde etkili deneylerdir. Araştırmacıların üç bağımsız değişkenin her birinin temel etkilerini belirlemesine olanak sağlar.

Karmaşık bir desende iki bağımsız değişkenin etkisi araştırıldığı zaman, sistematik değişimin potansiyel üç kaynağı yorumlanabilir. Her bir bağımsız değişken istatistiksel olarak anlamlı bir temel etki doğurabilir ve iki bağımsız değişken arasında anlamlı bir etkileşim ortaya çıkabilir. Betimleyici istatistikler bir tabloda veya bir çizgi grafikte paralel olmayan çizgilerle gösterildiği zaman, etkileşim ilkin çıkarma yöntemiyle belirlenebilir. Eğer etkileşimin kesinlikle anlamlı olduğu gösterilirse, basit temel etkileri ve eğer zorunlu ise, bir kerede iki kez karşılaştırılması düşünülen ortalamaları daha ayrıntılı inceleyerek sonuçları analiz edebiliriz. Hiçbir etkileşim etkisi ortaya çıkmazsa, bağımsız değişkenin temel etkisini inceleriz ve gerekliyse, iki ortalamanın karşılaştırılmasını ya da güven aralıklarını kullanabiliriz.

Karmaşık desenler psikoloji teorilerinden üretilmiş kestirimlerin test edilmesinde önemli bir rol oynar. Karmaşık desenler aynı zamanda, teoriler test edilirken ortaya çıkan çelişkilerin çözümlenmesinde de önemlidir. Karmaşık bir desen kullanılırsa ve hiçbir etkileşim ortaya çıkmazsa, her bir bağımsız değişkenin etkisinin diğer bağımsız değişkenin veya değişkenlerin düzeylerine genellenebileceğini biliriz. Bir etkileşim etkisi ortaya çıktığı zaman, yine de bir bulgunun dış geçerliğinin sınırları açık bir şekilde belirlenebilir. Etkileşimin

olması için şu gereklidir: Davranışı doğrudan etkileyen (temel etkileri doğuran) değişkenleri ve diğer bağımsız değişkenle birlikte araştırıldığı zaman etkileşim doğuracak değişkenleri içerecek ilişkili bir bağımsız değişken belirleriz. Tavan ve taban etkileri gibi ölçme sorunları yüzünden ortaya çıkabilen etkileşim, iki bağımsız değişkenin gerçek ortak etkisini gösteren etkileşimle karıştırılmamalıdır. Etkileşim, aynı zamanda doğal gruplar deseninde nedensel çıkarımlar elde etme probleminin çözümünde de çok yararlıdır.

TEMEL KAVRAMLAR

karmaşık desenler 244
temel etki 247
etkileşim etkisi 250
basit temel etki 259
ilişkili bağımsız değişken 265
tavan ve taban etkileri 266

GÖZDEN GEÇİRME SORULARI

1 Karmaşık desen deneylerinin aşağıdaki örneklerinin her birinde, bağımsız değişkenlerin sayısını, her bağımsız değişken için düzey sayısını ve toplam koşul sayısını belirleyiniz: (a) 2 x 3 (b) 3 x 3 (c) 2 x 2 x 3 (d) 4 x 3.
2 Şu bağımsız değişkenler faktöryel olarak birleştirildiğinde, karmaşık bir desendeki koşulları söyleyin: (1) Üç düzeyi olan görev türü (görsel, işitsel, dokunsal) ve (2) İki düzeyde test edilen çocuklar grubu (gelişimsel geriliği olanlar/olmayanlar).
3 Etkileşim etkisini belirlemede kullanılan iki yolu göstermek için Kassin ve arkadaşlarının Tablo 8.3'teki sorgulayıcıların itiraf elde etme çabalarına ilişkin sonuçlarını kullanın.
4 2 x 2'li karmaşık bir desenin sonuçlarını gösteren bir tabloda, etkileşim etkisi olup olmadığına karar vermek için, çıkarma yöntemini nasıl kullanacağınızı açıklayın.
5 Karmaşık bir desende etkileşimin ortaya çıktığını gösteren bir çizgi grafik örüntüsünü açıklayın.
6 İki bağımsız değişkenli karmaşık bir desende etkileşim etkisi olduğunda ve olmadığında yapılacak analizin basamaklarını özetleyin.
7 Bir psikoloji teorisinden üretilmiş kestirimleri test ederken, karmaşık bir desenin nasıl kullanılabileceğini gösteren bir örnek verin.
8 Karmaşık bir desende etkileşimin olması veya olmaması, bulguların dış geçerliğini nasıl etkiler?
9 Bir bağımsız değişkenin davranış üzerinde bir etkisi olmadığını söylerken, araştırmacıların niçin dikkatli olmaları gerektiğini açıklayın.
10 Bir veri setinde bir tavan (ya da taban) etkisi olabileceğini gösteren betimleyici istatistiksel örüntüyü açıklayın ve bu veri örüntüsünün çıkarımsal istatistiğin yorumunu nasıl etkileyebileceğini açıklayın.
11 Karmaşık bir desendeki etkileşimin, doğal gruplar deseninde nedensel çıkarımlara ulaşma sorununu çözmek için nasıl kullanılabileceğini açıklayın.

DÜŞÜNDÜRÜCÜ SORULAR

1 İki bağımsız değişkenin değişimlendiğini düşünün. Değişken *A* üç düzeyde ve değişken *B* iki düzeyde değişimlenmiştir.
- **A** Değişken *B*'nin temel etkisi olduğunu, değişken A'nın temel etkisinin olmadığını ve iki değişken arasında bir etkileşim olmadığını gösteren bir grafik çizin.
- **B** Değişken *A* ve *B*'nin temel etkisi olmadığını, fakat ikisi arasında bir etkileşim olduğunu gösteren bir grafik çizin.
- **C** Değişken *A* ve *B*'nin temel etkisi olduğunu ve ikisi arasında etkileşim olmadığı durumu gösteren bir grafik çizin.

2 Araştırmacılar bir deneyde şüphelilerin haklarından vazgeçmelerine yol açan etkenleri incelemek için, Miranda'nın uyarısını duyduktan sonra öğrencilerin haksız yere kopya çekmekle suçlandıkları bir durum yarattılar ve bu dersin hocası ile bir araya geldiklerinde, bir avukat isteme haklarından vazgeçmek isteyip istemediklerini sordular (bk. Scherr & Madon, 2013). Yaygın bir polis taktiği, Miranda'nın haklarını önemsizleştirerek (örneğin, hakların önemli olmadığını söyleyerek) suçlanan bireylerin haklarından vazgeçmelerini sağlamaktır. Bu nedenle araştırmacılar hakkından vazgeçmeyi iki düzeyde değişimlediler. Öğrencilere haklarından vazgeçmelerine ilişkin kararın sonucunun, durumları açısından önemli ya da önemsiz olduğu söylendi. Araştırmacılar aynı zamanda stresin şüphelilerin haklarını anlamalarını etkilediği hipotezini de kurdular. Öğrencilere kopya çekmelerinin büyük ya da küçük bir ihlâl olduğunu söyleyerek stresi değişimlediler. Yirmi öğrenci seçkisiz olarak deneyin her bir koşuluna atandı. Araştırmacılar öğrencilerin bir avukat bulundurma haklarına ilişkin anlayışlarını ölçtüler.
- **A** Bağımsız değişkenleri ve faktöryel birleşimin kullanıldığı durumdaki deney koşullarını da içerecek şekilde, bu çalışmanın desenini betimleyin.
- **B** Anlama, katılımcılardan "Söylediğiniz herhangi bir şey size karşı kullanılabilir ya da kullanılacaktır." gibi beş cümlenin açıklaması istenerek değerlendirildi. Açıklamalar 0 (kötü) ile 10 (iyi) arasında değişen olası anlama puanları olarak kaydedildi. Bu ölçüme ilişkin ortalamalar aşağıdadır. Bağımsız değişkenlerin herhangi bir etkisini (sıfırdan daha büyük ortalama farklılıklarının anlamlı olduğunu varsayın) betimleyin.

Hakından Vazgeçme Tanımı	Stres	
	Ağır	Küçük
Önemsiz	7.0	8.0
Önemli	7.6	8.6

3 Bir deneydeki katılımcıların bir diğer katılımcıya yardım edip etmeyeceğini test etmek için, bir araştırmacı ekranın en altında bulunan bir robot karakterin, ekranın her yönünden seçkisiz olarak gelen topları yakalamak üzere manevra yaptığı bir bilgisayar oyunu geliştirdi. Ekran ortadan ikiye bölündü. Katılımcılara ekranın bir yarısının kendi oyunlarını, diğer yarısının ise yan odada oynayan bir başkasının oyununu gösterdiği söylendi (gerçekte, yanda kimse yok). Katılımcılar klavyeye dokunarak, kendi oyunu ile diğer kişinin oyununu gösteren iki görüntü arasında gidip gelebilecekti, fakat diğer oyuncu onların robotunu göremeyecekti. Diğer oyuncunun yakaladığı herhangi bir top (oyundaki diğer toplar gibi) basitçe ekranda kaybolacaktı. Oyunun ortasında, diğer oyuncunun robotu ekranın bir tarafında takılı kalmış gibi göründü. Bağımlı değişken olarak yardım etme ölçütü, katılımcının diğer oyuncu için ne kadar top yakaladığı idi.

Araştırmacı, katılımcıların dikkatleri kendi üzerlerindeyken, yardım etmelerinin daha az olası olduğu hipotezini kurdu. Temposu hızlı bir oyunun, katılımcıların daha fazla kendi oyunları üzerine odaklanmalarına yol açacağı sonucuna vardı; bu yüzden katılımcı oyun temposunu hızlı ve yavaş olmak üzere iki düzeyde değişimledi. Araştırmacı, puanlarını yükseltmeleri için katılımcılara ödül verilirse, yardım etme olasılıklarının daha az olacağı kestirimini de yaptı (100 toptan fazla yakalandığında para verilmesi). Ödülü üç düzeyde değişimledi: Sıfır ödül, 5 $ ve 10 $. Yirmi katılımcı, bu karmaşık desenin her koşulunda test edildi. Her koşulda diğer oyuncu için yakalanan top sayısının ortalaması aşağıda verilmiştir:

Oyun Hızı	Ödül		
	Sıfır	5$	10$
Yavaş	14	10	4
Hızlı	4	2	1

- **A** Bu deneyde bir etkileşim olabileceğine dair bir kanıt var mıdır?
- **B** Eğer deney odasında katılımcıdan başka birisi daha olsaydı araştırmanın hangi yönleri sizi bir etkileşimi yorumlarken daha temkinli olmaya sevk ederdi?
- **C** Araştırmacı bir etkileşimi test edebilmek için deneyinde nasıl değişiklikler yapabilirdi?

(devam ediyor)

4 Bir psikolog yaşlı insanların karmaşık şemaları daha yavaş işledikleri (processing) hipotezini kurdu. İçinde saklı şekiller bulunan bir testi kullanarak, 20 yaşlı (65-70 yaş) ile 20 genci (18-23 yaş) test etti. Uyguladığı işlemde, bir tablet ekranında önce basit bir şekil göründü; sonra bu basit şekli de içeren karmaşık şekil göründü. Katılımcının görevi 10 denemenin her birinde olabildiğince hızlı bir şekilde basit şeklin, daha karmaşık olan ikinci şekil içindeki yerini belirlemekti. Katılımcı şekli bulduğunda, tablette şekli bulduğu yere dokundu. Sonuçlar genç gruba kıyasla yaşlı grubun, basit şekli daha geç bulduğunu gösterdi. Fark istatistiksel olarak anlamlıydı.

A Psikolog tepki zamanı farkının, yaşlılar arasında azalan karmaşık bilgiyi işleme yeteneğindeki azalmadan kaynaklandığı sonucuna vardı. Bu özetteki hangi bilgi, psikoloğun bu nedensel çıkarımı yapmasına olanak tanır? Hangi bilgi eksiktir?

B Bir diğer araştırmacı, gençlerin hızı, yaşlıların ise hız yerine doğruluğu tercih edeceği hipotezini kurdu (örneğin, tepki vermeden önce doğru cevapladıklarından emin olmak isterler). Araştırmacı desene iki düzeyi olan bir başka değişken (ödül) ekledi. "Doğruluk" koşulundaki katılımcılar doğru sonuçları için, "hız" koşulundaki katılımcılar ise tepki hızları için ödüllendirildiler. Bu deneyin desenini, deneydeki koşulları içerecek şekilde betimleyin.

C Bağımlı değişken, doğru belirlemelere verilen ortalama tepki zamanıydı (saniye olarak). Aşağıdaki verilerde hangi etkiler vardır? (Katılımcı sayılarının eşit olduğunu ve sıfır olmayan farkların istatistiksel olarak anlamlı olduğunu varsayın). Sonuçlar ikinci araştırmacının teorisi hakkında ne söylemektedir?

	Ödül	
Yaş Grubu	Doğruluk	Hız
Genç	16.0	10.0
Yaşlı	24.0	12.0

İleri Alıştırma I'in Cevapları

1 a) Gerçek Suçluluk: $M = 3.04$, Gerçek Suçsuzluk: $M = 3.18$

b) Şüpheli statüsü değişkenine ait ortalamalar arasındaki fark 0.14'tür. Bu fark, sorgulayıcı beklentisinin, varsayımsal soru sayısı üzerinde anlamlı olduğu gözlenen etkisine ait ortalama farkla (3.62 -2.60 =1.02) kıyaslandığında, oldukça küçük bir farktır.

c) Çıkarma yöntemi kullanarak, suçlu beklentisi koşulunda, gerçek suçluluk ve gerçek suçsuzluk koşulları arasındaki farkın - 0.16 (3.54 -3.70) olduğu görülür. Suçsuzluk beklentisi koşulunda bu fark – 0.12'dir (2.54 – 2.66). Bu farklar çok benzer olduğu için, bir etkileşim olması mümkün değildir.

2 (a) Suçlu beklentisi: $M = 9.84$, Suçsuz beklentisi: $M = 8.74$

(b) Suçlu beklentisi ve suçsuz beklentisi koşullarının ortalamaları arasındaki fark 1.1'dir.

(yani, suçlu beklentisi koşulundaki ikna edici teknik sayısı, suçsuz beklentisi koşulundakinden yaklaşık bir fazladır). Tersine, şüpheli statüsü bağımsız değişkeni temel etkisine ait gerçek-suç ve gerçek-suçsuz koşullarında kullanılan ikna edici teknikler arasındaki fark 4.27'dir (11.42-7.15).

c) Bir etkileşim etkisi olası değildir. Çıkarma yöntemini kullanarak, suçlu beklentisi koşulunda, gerçek suçlu ve gerçek suçsuz koşulları arasındaki fark (7.71 − 11.96 = −4.25), suçsuz beklentisi koşulundaki değere (6.59 − 10.88 = −4.29) çok benzerdir.

3 (a) Suçlu beklentisi: $M = 6.40$, Suçsuz beklentisi: $M = 5.70$

(b) Gerçek suçluluk: $M = 5.60$, Gerçek suçsuzluk: $M = 6.51$.

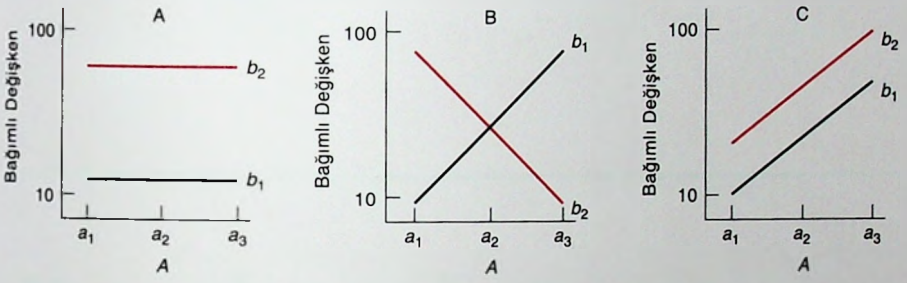
(c) İstatistiksel olarak anlamlı sorgulayıcı-beklentisi değişkeninin temel etkisi, itiraf alma çabasının (bağımlı değişken), suçlu beklentisi koşulunda ($M = 6.40$), suçsuz beklentisi koşulundan ($M = 5.70$) yüksek olduğunu gösteriyor.

İstatistiksel olarak anlamlı şüpheli statüsü değişkeninin temel etkisi gösteriyor ki gerçek suçsuz koşulundaki itiraf elde etme çabası ($M = 6.51$), gerçek-suçlu koşulundakinden daha yüksektir ($M = 5.60$).

İleri Alıştırma II'nin Cevapları

A Etkileşim etkisi, görev zorluğunun temel etkisi
B Etkileşim etkisi yok, görev zorluğu ve kaygı düzeyinin temel etkileri
C Etkileşim etkisi yok, medya ve içerik tipinin temel etkileri
D Etkileşim etkisi, medya ve içerik tipinin temel etkileri
E Etkileşim etkisi, gecikme ve örüntü karmaşıklığının temel etkileri (bu etkileri test etmek için ek istatistiksel analizler gerekir)
F Etkileşim etkisi yok, gecikme ve örüntü karmaşıklığının temel etkileri

Düşündürücü Soruların Cevapları



KISIM DÖRT

Uygulamalı Araştırma

BÖLÜM DOKUZ

Tek Denekli Araştırma Desenleri

BÖLÜMÜN ANA HATLARI

GENEL BAKIŞ
VAKA ÇALIŞMASI YÖNTEMİ
Özellikleri
Vaka Çalışması Yönteminin Avantajları
Vaka Çalışması Yönteminin Dezavantajları
Vaka Çalışmasına Dayanan Tanıklıklar Üzerine Eleştirel Düşünme
TEK DENEKLİ (KÜÇÜK N) DENEYSEL DESENLER
Tek Denekli Deneylerin Özellikleri
Özel Deneysel Desenler
Tüm Tek Denekli Desenlerin Ortak Sorunları ve Sınırlılıkları
ÖZET

GENEL BAKIŞ

Bu bölüme kadar kitabınızda, bir veya birden fazla katılımcı gruplarının ortalama performanslarını incelemek için tasarlanan *grup yöntemleri* üzerinde durduk. 6, 7 ve 8. bölümlerde deneysel yöntemleri anlatırken, bu durum özellikle belirgindi. Bu bölümde ise, tek bir bireyin incelendiği iki alternatif yöntemi tanıtacağız. Bu yöntemler *tek denekli araştırma* desenleri olarak adlandırılırlar.

Tek denekli desenler 19'uncu yüzyılda bilimsel psikolojinin başlangıcından itibaren kullanılmaktadır. Psikofizik yöntemler, Gustav Fechner'in çalışmalarıyla başlamış ve 1860'ta yazdığı "*Elemente der Psychophysik*" adlı kitabında anlatılmıştır. Fechner ve sayısız psikofizikçi, çalışmalarını bir veya iki bireyden elde ettikleri verilere dayandırmıştır. Hermann Ebbinghaus da tek-denekli deseni kullanmış erken psikoloji tarihinin önemli isimlerinden biridir. Aslında Ebbinghaus'un incelediği tek kişi de kendisi olmuştur. 1885'te bellek hakkında yayımladığı monografında (tekyazı), araştırmanın hem katılımcısı hem de deneycisi olmuştur. Birkaç ay süresince yüzlerce anlamsız hece dizisini öğrenmiş ve sonra tekrar öğrenmeye çalışmıştır. Onun verileri psikologlara, unutmanın zaman içinde nasıl meydana geldiğini gösteren ilk sistematik kanıtları sağlamıştır.

Vietnam gazileri için bilişsel terapiden (Kubany, 1997), amnezi hastalarının beyin süreçlerinin incelenmesine (Gabrieli, Fleischman, Keane, Reminger ve Morrell, 1995) ve Tourette sendromu ile ilişkili motor ve vokal tiklerin tedavisine (Gilman, Conor ve Haney, 2005) kadar çok farklı konuları ele alan tek denekli çalışmalar, düzenli olarak psikoloji dergilerinde yayımlanmaktadır. Balerin, satranç oyuncusu veya müzisyen gibi uzmanların performansını inceleyen bilişsel psikologlar da bu yöntemleri kullanmaktadırlar (örneğin, Ericsson ve Charness, 1994). Örneğin, araştırmacılar "bugüne kadar anlatılan en hızlı ve doğru takvim dâhisi olan otistik bilgin Donny" ile ilgili gözlemlerini rapor etmişleridir (Thioux, Stark, Klaiman ve Schultz, 2006, s. 1155). Bir saniyeden az bir sürede haftanın hangi gününde doğduğunuzu söyleyebilen Donny'e altı yaşındayken otizm tanısı konmuş ve sınır zekâ IQ puanına sahip olduğu bulunmuştur. Buna rağmen 1 Mart 1990 ile 28 Şubat 2100 tarihleri arasında test edildiğinde %98 doğru cevap verdiği görülmüştür.

Bu bölümde, vaka çalışması yöntemi ve tek denekli deneysel desenler olmak üzere iki özel tek-denekli araştırma yöntemini tartışacağız: Vaka çalışması sıklıkla klinik psikoloji alanı ile ilişkilendirilir, ancak antropoloji, kriminoloji, nöroloji ve sosyoloji gibi alanlarda çalışan araştırmacılar da bu önemli yöntemi kullanırlar. Örneğin, Oliver Sacks (2007, 2012) adlı nörolog, olağandışı ve etkileyici beyin bozuklukları olan kişilere ilişkin vaka çalışmaları ile milyonları büyülemiştir. "*Halüsinasyonlar*" adlı kitabında Sacks (2012) var olmayanı deneyimlemenin, gerçek olmayanı koklamanın, duymanın ve görmenin nasıl bir şey olduğunu tarif etmektedir. Sacks burada psikotik bir durumla (örneğin, şizofreni) ilişkili bir sanrıdan değil, sıra dışı olmayan, hepimizin yatkın olduğu sanrı çeşitlerinden bahsetmektedir. Bunlar ilaç alımı, migren baş ağrıları, duyusal yoksunluklar (örneğin, körlük) veya monotonluk (örneğin, değişmeyen peyzaj veya deniz manzarasına sürekli bakmak) nedeniyle ortaya çıkabilirler. Sanat ve edebiyatta halüsinasyonların oynadığı rolü incelemiş, örneğin, *Alice*

Harikalar Diyarında kitabının yazarı Lewis Caroll'un migren sancılarına neden olabilecek baş ağrılarından şikâyet ettiğine işaret etmiştir. Sack'ın deyimiyle bu "klinik hikayeler" sadece beyin ve zihin arasındaki ilişkiye iç görü kazandırmakla kalmamakta, insan olmanın ne anlama geldiğine ilişkin anlayışımızı da genişletmektedir. Vaka çalışması yönteminin güçlü yönlerini ve sınırlılıklarını daha sonra ele alacağız.

Vaka çalışmasının aksine, tek denekli deneysel desende tipik olarak değişkenlerin değişimlenmesi (manipulation) ve tek bir denek için yorumlanması öne çıkmaktadır (birkaç denek veya tek bir grup gözlense bile). Tek-denekli deneysel desenler sıklıkla "N = 1 deneysel desenleri" veya " küçük-*n* araştırma desenleri" olarak anılırlar. Bu desenler, *davranışın deneysel analizi* ve *uygulamalı davranış analizi yaklaşımlarının* örnekleridir. Göreceğiniz üzere, bu yaklaşımlar küçük-*n* yaklaşımının sırasıyla temel ve uygulamalı yaklaşımını temsil ederler. Tek denekli desenler vaka çalışmalarından daha sistematik ve kontrollüdür. Bu desenlerin altında yatan muhakemeyi inceleyecek ve sık kullanılan tek denekli deneysel desenlere örnekler vereceğiz. Bu deneysel desenler, 7'nci Bölüm'de anlatılan tekrarlı ölçümler deseninin özel bir durumunu temsil etmektedir.

VAKA ÇALIŞMASI YÖNTEMİ

Özellikleri

- Bireylerin ayrıntılı şekilde betimlenmelerini ve analiz edilmelerini içeren vaka çalışmaları, küçük-*n* araştırma desenlerinde bulunan kontrolden yoksundur.
- Vaka çalışmaları, normal ve anormal davranışa ilişkin düşünce ve hipotezler geliştirmek için kaynak oluşturur.

Temel Kavram

Vaka çalışması, tek bir bireyin yoğun biçimde betimlenmesi ve analizidir. Vaka çalışmaları, her zaman olmasa da sıklıkla niteliksel veriden yararlanır (örneğin, Smith, Harré ve Van Langenhove, 1995). Bu yöntemi kullanan araştırmacılar doğal gözlem, arşiv kayıtları (4'üncü Bölüm), görüşme ve psikolojik testleri (5'inci Bölüm) içeren birçok kaynaktan veri toplar. Örneğin, Afrika'da vahşi hayvan ve evcil hayvan ticaretinden kurtarılmış olan iki şempanze, Poco ve Safari'nin vaka çalışmalarını ele alalım (Lopresti-Goodman, Kameka, ve Dube, 2013). Bu şempanzeler barınaklarda "yaşamlarının ilk 7-8 yılını yalnız ve insanlara sergilenerek geçirmişlerdir" (s. 1). Araştırmacılar bakıcılarla görüşmelerden ve doğrudan gözlemlerden elde edilen bilgileri birleştirerek, psikolojik hasarı değerlendirmişlerdir. Ne yazık ki kurtarılmalarının ardından, erken travmadan kaynaklanan stereotipik davranışlarını onlarca yıl sürdürdükleri görülmüştür (bk. Kutu 9.1).

Klinik bir vaka çalışması çoğu zaman belirli bir tedavinin uygulamasını ve bulgularını açıklar. Örneğin, klinik bir vaka çalışması bireyin belirtilerini, bu belirtileri anlamak ve tedavi etmek için kullanılan yöntemleri ve tedavinin etkililiğine ilişkin kanıtları betimleyebilir. Klinik vaka çalışmalarında kullanılan

KUTU 9.1

ŞEMPANZELERDE TRAVMA SONRASI STRES BOZUKLUĞU

Vahşi doğada yavru şempanzeler yaşamlarının ilk beş yılını anneleriyle geçirirler. Ancak çeşitli Afrika ülkelerinde yetişkin şempanzeler, etleri için kanunsuz olarak avlanır ve yavruları evcil hayvan olarak satılır. Sonuç olarak bu yetim şempanzeler anneden erken ayrılığı ve yıllarca kısıtlı tutsaklık deneyimini yaşarlar (örneğin, turistleri çekmek için kafeslerde tutulurlar). Kenya Nanyuki'de bulunan Sweetwater Şempanze Barınağı, yetim kalmış ve/veya el konulmuş şempanzelere ayrılmış birçok barınaktan biridir (bk. Lopresti-Goodman ve ark., 2013).

Barınaktaki araştırmacılar bir vaka çalışması hazırlayarak, yetişkin şempanzelerin sıkıntılarını değerlendirmeyi amaçlamışlardır. Gashuhe, 9 yaşında kurtarılmış olan 14 yaşındaki yaşlı bir erkek şempanzedir. Bu zamana kadar dışarıdaki etmenlerden bağımsız, kendine ait bir küçük açık hava kafesinde tek başına yaşamıştır. Araştırmacıların özellikle ilgi noktası, Gashuhe'nin anormal davranışlarının (örneğin, dışkı yeme, sallanma, kendini kenetleme, kendini yaralama gibi) insanlardaki psikopatolojiye ne kadar benzer olduğudur. Araştırmacılar bakıcılarla derinlemesine görüşmeler yapmışlar ve 7 gün boyunca Gashuha'nın davranışlarını gözlemişlerdir. Gözlemciler, insanlardaki psikopatolojiye ilişkin belirtileri kapsayan bir listenin şempanzelere uyarlanmış bir versiyonunu kullanmışlardır (bk. Ferdowsian ve ark., 2011).

Gashuhe'nin DSM-5'te betimlenen bir kaygı bozukluğu olan travma sonrası stres bozukluğu kriterlerine uyduğu gözlenmiştir: Kolayca kaygılanma veya uyarılma, bazı yerlerden ve insanlardan kaçınma, aşırı patlamalar ve sürekli tetikte bekleme (hypervigilans). Buna ek olarak, vücudunun bazı parçalarının ona yabancı geldiği ve ona saldırdığını deneyimlediği çözülme anları yaşadığı, buna karşı mücadele etmek için saldıran bacağını ısırdığı veya güreştiği gözlenmiştir. Şempanzelere psikopatoloji teşhisi konulması tartışılır olsa da bu vaka çalışması, şempanzeleri evcil hayvan olarak alıkoymanın yaşam boyu yıkıcı sonuçlarına dikkat çekmekte ve onlara da insana özgü zihinsel hastalıklar için başvurulan tedaviler önermektedir (DSM-5, 2'nci Bölümde tartışılmıştır. Gashuhe ile ilgili bilgi Dr. Lopresti-Goodman tarafından sağlanmıştır).

ŞEKİL 9.1 (a) Gashuhe travma ilişkili bozukluğu bulunan yetim bir şempanzedir. (b) Dr. Lopresti-Goodman gibi davranış bilimciler vaka çalışması yöntemini, erken yaşta anneden ayrılmanın ve sosyal izolasyonun şempanzelerdeki uzun süreli etkilerini değerlendirmek için kullandılar.





işlem değişkenleri nadiren sistematik olarak kontrol edilir. Buna karşılık, birçok işlem aynı anda uygulanabilir ve psikologların dış değişkenler (örneğin, hastanın belirtilerini etkileyebilecek iş veya ev ortamları) üzerinde çok az kontrolü olabilir. Sonuç olarak, *vaka analizlerinin temel özelliklerinden biri yüksek düzey kontrolden yoksun olmalarıdır.* Kontrolün olmadığı durumda araştırmacıların, bireyin davranışlarını (tedavi de dahil) etkileyen değişkenlere ilişkin geçerli çıkarımlar yapmaları da kolay değildir. Kontrol düzeyi, vaka çalışması ile tek denekli deneysel desenler arasındaki ayırıcı bir özelliktir ve tek denekli desenlerde kontrol düzeyi daha yüksektir.

Vaka çalışması yöntemi genellikle "keşif amaçlı" olarak nitelendirilmekte ve davranış hakkında düşünce ve hipotezlere kaynak oluşturmaktadır (Bolgar, 1965). Diğer yandan, deneysel yaklaşımlar ise belirli hipotezleri test etme fırsatı olarak görülür. Vaka çalışması yöntemi bazen daha kontrollü inceleme yöntemlerine karşıt gibi görülmüşse de daha doğru bir bakış açısı, *vaka çalışmalarının psikolojideki diğer yöntemlerle ilişkili ve tamamlayıcı olduğudur* (Kazdin, 2011).

Vaka çalışması yönteminin araştırmacı psikolog için hem avantajları, hem de dezavantajları vardır. Ama avantaj ve dezavantajlarını anlatmadan önce Kirsch'ün (1978) bildirdiği gerçek bir vakanın özetini sunarak bu yönteme bir örnek vereceğiz. Daha sonra avantaj ve dezavantajlarına ilişkin yapacağımız tartışmada bu örneği kullanacağımız için bu kısaltılmış vaka çalışmasını dikkatlice okumanız önemlidir (Kutu 9.2).

KUTU 9.2

HASTALAR KENDİ KENDİLERİNİN TERAPİSTİ OLABİLİRLER Mİ? BİR VAKA ÇALIŞMASI ÖRNEĞİ

Bu makale kısa terapilerin avantajlarından faydalanan, ama aynı zamanda birçok görevin eksik bırakılmasından doğan sakıncaları da azaltan bir terapi stratejisi olan öz yönetim eğitiminin (ÖYE) kullanımı üzerinedir. Bu yaklaşımın özü, hastaya kendi kendinin terapisti olmayı öğretmektir. Hastaya, problemleri davranışsal boyutlarda değerlendirmesi ve problemlerin üstesinden gelmek için mevcut tedavi tekniklerine dayanarak belirli taktikleri geliştirmesi öğretilir. Hasta, hasta ve terapist ikili rolünü üstlenirken, terapist de danışmanlık görevini üstlenir.

Susan vakası

28 yaşında evli bir kadın olan Susan, bellek bozukluğu, düşük zekâ ve özgüven eksikliği şikâyetleriyle terapiye gelmişti. Susan'ın var olduğuna inandığı bu sorunlar, onun bazı sosyal durumlarda ketlenmesine "neden" olmaktaydı. Susan film, tiyatro oyunları, kitap ve dergiler ile ilgili tartışmalara, onları yeterince iyi hatırlamadığı "için" girememekteydi. Genellikle bir konuşmada nelerin anlatıldığını anlayamadığını düşünmekte, bunun da onun zekâ düşüklüğünden kaynaklandığı sonucuna varmaktaydı.

Bu etkileşimlerde pasif bir rol takınarak anlama güçlüğünü saklamaya çalışmakta, sorulacak sorularla bu yetersizliğinin keşfedileceğinden endişe etmekteydi. Kendi fikrine güvenmemekte ve hatta bazen bir fikri olduğundan bile şüphe etmekteydi. Bir fikir edinmek için başkalarına bağımlı olduğunu hissetmekteydi.

Wechsler Zekâ Testini (WAIS) uyguladım ve hiç de normalin altında olmayan, 120 sözel zekâ puanı olduğunu buldum. Rakam uzamı puanı en azından kısa süreli belleğinin bozuk olmadığını gösterdi. Test, onunla konuşmalarım sırasında da tahmin ettiğim gibi, zekâ düzeyinde ve belleğinde bir sorun olmadığını gösterdi. Bu sonucu tartıştıktan sonra, eğer belleğinin, zekâsının ve özgüveninin yeterli olduğunu düşünseydi, neler yapabileceğini ayrıntılı şekilde incelemeyi önerdim. Bu yolla, bir kanaat belirtme, açıklama isteme ve bazı bilgileri bilmediğini itiraf etme gibi bazı davranışsal hedeflerde anlaşma sağladık. Terapi seanslarında Susan'a, kaygı yaratan ortamlarda açık ve kapalı provalar yapması gibi yapılandırılmış ev ödevleri yapması için rehberlik ettim. Bu ödevler, davranışsal hedeflerine adım adım yaklaşmayı ve gelişimini kaydetmeyi içermekteydi. Buna ek olarak, test sonuçları doğrulamadığı hâlde kendisine atfettiği olumsuz ifadeleri (örneğin, "Ben aptalım.") tartıştık. Böyle bir olumsuz cümle kullandığını fark ettiğinde, bilinçli olarak bunu daha olumlu bir ifade ile değiştirmesini önerdim (örneğin, "Aptal değilim - aptal olduğumu düşündürecek mantıklı bir neden yok.").

Terapinin beşinci seansında Susan, zor olduğunu düşündüğü bir ev ödevini tamamladığını bildirdi. Yapmayı sadece kolay bulmakla kalmamış, birinci denemede bile endişe duymadığını bildirmişti. İşte tam bu noktada, terapi ilişkisinin niteliği değiştirildi. Sonraki seanslarda Susan hafta içindeki performansını değerlendirdi, bir sonraki aşamanın ne olacağını belirledi ve kendi ev ödevlerini hazırladı. Benim rolüm bir terapi öğrencisine danışmanlık yapmaya dönüştü; başarısını pekiştirdim, göz ardı ediyor olabileceği ögelere dikkat çektim.

Dokuzuncu terapi seansından sonra doğrudan terapiye devam edilmedi. Sonraki ay içerisinde iki kez Susan'ı telefonla aradım. Hedeflerine ulaşmada kendine güven duyduğunu bildirdi. Özellikle, hayatında yeni bir kontrol algısı hissettiğini belirtti. Benim izlenimim, davranışsal bir problem çözme değerlendirme yöntemini başarıyla edindiği ve hedeflerini gerçekleştirmek için stratejiler geliştirmede oldukça beceri kazandığı yönündedir.

İzlem

Tedavinin sonlanmasından beş ay sonra, Susan'la iletişim kurarak gelişimi konusunda bilgi istedim. Sosyal ortamlarda eskisinden daha çok konuştuğunu, kendi başına (örneğin, eşi olmadan) bir şeyler yaparken kendini daha rahat hissettiğini ve genel olarak artık aptal olduğunu düşünmediğini bildirdi. "Önceden düşündüğüm yerden bir basamak veya bir düzey daha yukarıda olduğumu hissediyorum." diyerek özetledi durumunu.

Terapide kullandığımız etkinliklerden hangilerini kullanmaya devam ettiğini de sordum. Son olarak terapinin sonlanmasını izleyen beş ay içerisinde en az üç kez bir başka kişiye "Bunu anlayamadım, bana tekrar açıklar mısın?" dediğini bildirdi. Bu, daha önce "aptallığını" karşısındaki kişiye belli etmemek için yapmayı beceremediği bir davranıştı.

İzlem görüşmesinden üç ay sonra Susan'dan talep etmediğim bir mektup aldım (o sırada başka bir eyalete taşınmıştım). Bu mektupta bana, imgeleme alıştırmalarından birinde folklor dersine girmeyi ve kendini iyi hissetmeyi hedeflediğini hatırlatmış ve nihayet bunu başardığını yazmıştı.

Kaynak: Kirsch, I. (1978). Teaching clients to be their own therapists: A case study illustration. *Psychotherapy: Theory, Research, and Practice*, 15, 302–305 (İzin alınarak tekrar basılmıştır.).

Vaka Çalışması Yönteminin Avantajları

- Vaka çalışmaları yeni fikirler ve hipotezler sağlar, yeni klinik teknikler geliştirme fırsatı ve nadir olguları inceleme fırsatı verir.
- Tek bir vakanın davranışı, kuramsal ilkeler veya iddialarla çeliştiğinde, bilimsel kuramlar sorgulanabilir veya vaka çalışmalarından gelen kanıtlar kuramlara destek verebilir.
- İdiyografik araştırma (sadece bir kişiye özgü olanın belirlenmesi için bireylerin incelenmesi) nomotetik (tipik olanı tanımlamak için grupların incelenmesi) araştırmayı tamamlar.

Davranışa İlişkin Düşüncelerin Kaynakları Vaka çalışmaları bireyler hakkında zengin bilgi kaynakları ve kişilerin davranışlarının nedenlerine ilişkin iç görü sağlar. Bu iç görüler araştırma hipotezlerine dönüştürüldüğünde daha kontrollü araştırma yöntemleri kullanılarak test edilebilirler. Susan adındaki hasta ile başarılı bir psikoterapi gerçekleştirmiş olan terapist de tartışmasında, vaka çalışması yönteminin bu yönünü onaylamıştır. Görüşünü "Bu vaka çalışmasının bulguları deneme amaçlı görülmelidir. Bu tekniğin yararının, daha kontrollü araştırmalarla ortaya konulacağını umut ediyorum." şeklinde belirtmiştir (s.305). Vaka çalışması, bir araştırmacı tarafından az bilinen bir alana girmek için doğal bir başlangıç noktasıdır.

Klinik Yenilik Fırsatı Vaka çalışmaları yöntemi, yeni terapi tekniklerini veya var olan tekniklerin farklı uygulamalarını denemek için fırsat sağlar. Örneğin, psikoterapide öz yönetim eğitimi (self management training) kullanımı, tipik hasta-terapist ilişkisini değiştirdiği için klinik bir yenilik içermektedir. Öz yönetim eğitimi yöntemi hastalara kendi kendilerinin terapisti olmayı öğretmeyi, diğer bir deyişle problemleri tanımlamayı ve onlarla baş etmek için teknikler tasarlamayı esas alır. Hasta hem hasta hem terapist iken, terapist danışman olarak görev yapar.

Nadir Olguları Çalışma Yöntemi Vaka çalışması nadir olayları çalışmak için de yararlıdır. Bazı olaylar doğada o kadar nadir ortaya çıkarlar ki, onları ancak yoğun bir biçimde inceleyerek betimleyebiliriz. Örneğin, Oliver Sacks'ın kitaplarında anlatılan vakaların birçoğunda az rastlanan beyin bozuklukları bulunmaktadır. Bölümün başında anlattığımız otistik bilginler ve olağanüstü belleğe sahip kişiler de vaka çalışmalarının nadir olayları incelemekte nasıl kullanıldığına örnektir.

Kuramsal Varsayımlara Meydan Okuma Marslılar üç başlıdır şeklindeki bir kuram, güvenilir bir gözlemci sadece iki başlı bir Marslı gördüğünde hemen çökecektir. Vaka çalışması yöntemi, genellikle evrensel olarak kabul edilen bir ilkeyi veya genel önermeyi ihlal eden "karşıt bir örnek" sunarak bilimsel düşünceyi ileriye götürür (Kazdin, 2011). Konuşmayı üretebilme ve işleyebilme becerisinin tonlamayı anlama becerimize dayandığını (özellikle Çince gibi tona bağımlı dillerde) öne süren bir kuramı düşünelim. Konuşmanın tonlaması ile sesin yükselip alçalmasına ilişkin beceri ve bazı konuşmalardaki müzikal yön

ile müziği algılama arasında benzerlik bulunacağı düşünülebilir. Bu durumda böyle bir kuram, müziği duyamayan bir kişide normal konuşma algısı ve üretiminin olmasını nasıl açıklar?

Oliver Sacks (2007) müziği duyma yeteneği doğuştan olmayan (amuzi) bazı vakaları ilişkilendirmektedir. Örneğin, hiçbir zaman müziği, en azından bizlerin duyduğu gibi duymamış bir kadından bahsetmektedir. Bu kişi, ezgiler arasında ayırım yapamaz, bir müzikal notanın diğerinden alçak mı yüksek mi olduğunu söyleyemez. Müziğin ona nasıl duyulduğu sorulduğunda ise yere tencere tava fırlatma sesine benzediğini bildirir. Teşhisi ancak yetmiş yaşlarında konulmuş ve bu olağandışı bozukluğu olan başka kişilerle tanıştırılmıştır. Bununla birlikte o ve amüzisi olan diğer kişiler, normal konuşma algısı ve üretimi göstermektedirler. Açıkça görülüyor ki, dil becerisi ve müzik algısı arasında yakın bağ kuran bir kuram, bu vaka çalışmalarını temel alarak değişiklikler yapmaya ihtiyaç duyacaktır.

Psikoloji Kuramına Geçici Destek Vaka çalışmasından gelen bir kanıt, bir psikoloji kuramına geçici destek sağlayabilir. Vaka çalışması sonuçları belirli bir hipotez için *nihai* (kesin) bir kanıt sağlamak için kullanılmasa da, vaka çalışmasının bulguları bazen psikolojik kuramı destekleyecek önemli kanıtlar sunabilir.

Vaka çalışmalarının psikolojik kurama destek olabileceğine ilişkin bir örnek bellek literatüründen gelmektedir. 1960'larda Atkinson ve Shiffrin'in insan belleğine ilişkin ortaya attıkları model, alandaki çalışmaları onlarca yıl etkilemiştir. Bilgi işleme süreçlerine temellendirilen bu model, kısa süreli bellek (KSB) ve uzun süreli bellek (USB) sistemlerini tanımlamıştır. Her ne kadar sayısız deney bulgusu belleğin bu ikili doğasına kanıt sağladıysa da Atkinson ve Shiffrin birkaç vaka çalışması sonucunu "bellek sistemindeki ikiliğin (dikotominin) belki de en ikna edici gösterimi" (1968, s. 67) olarak kabul etmişlerdir. Bu vaka çalışmaları, epilepsi tedavisi için beyinlerinin temporal lobundaki bazı bölümleri (hipokampus denilen subkortikal alan da dahil) cerrahi olarak çıkarılan hastaları içermekteydi. Araştırmacıların kuramı için en ilginç vaka çalışması H.M olarak bilinen hastaydı (bk. Hilts, 1995; Scoville ve Milner, 1957). Beyin ameliyatından sonra H.M.'nin ciddi bellek hasarı olduğu bulundu. Normal sohbeti yürütebiliyor ve olayları kısa bir süre aklında tutabiliyordu ancak bir günden diğerine olayları hatırlayamıyordu. Aynı dergiyi defalarca okuduğu halde içeriğini tanıdık bulmazdı. H.M.'in kısa süreli belleği bozulmamış görünüyordu ancak bilgileri daha uzun süreli bir sisteme aktaramıyordu. H.M.'nin ve benzer bellek bozukluğu olan hastaların daha sonraki testleri, sorunun başlangıçta düşünülenden daha karmaşık olduğunu gösterdi. Buna rağmen H.M. vakası insan bellek kuramlarının tartışıldığı yerlerde önemini korumaya devam etmektedir (örneğin, bk. Schacter, 1996 ve Kutu 9.3).

Nomotetik Yöntemi Tamamlayıcı Psikoloji (genel olarak tüm bilimler gibi) geniş organizma popülasyonları için geçerli olacak "evrensel ilkeler", kapsamlı genellemeler yapmaya çalışmaktadır. Bunun bir sonucu olarak psikoloji araştırmaları, genellikle nomotetik yaklaşımı kullanan araştırmalar olarak nitelendirilir.

KUTU 9.3

PSİKOLOJİYE IŞIK VERMEYE DEVAM EDEN BİR TEK VAKA

Psikoloji araştırmacılarının elli yıldan fazla bir zaman H.M. olarak bildiği Henry Gustav Molaison, 2 Aralık 2008 yılında aramızdan ayrıldı. 82 yaşındaydı ve hayatının büyük kısmında sadece "bulunduğu an"da yaşadı; birkaç dakikadan fazlasının ve psikolojiye yaptığı katkıların farkında olmadı (metne bakınız). Los Angeles Times gazetesinde yayımlanan ölüm ilanında (T.H. Maugh, II, 9 Aralık, 2008), Nobel ödüllü Eric Kandel, "Bir vaka bütün bir bilgi alanını aydınlattı" ifadesini kullandı. Ancak H.M.'nin bilime katkısı ölümüyle de sona ermedi. Yıllar önce bir akrabasına danışarak beynini öldükten sonra bilime bağışlamayı kabul etti (bk. *B. Carey, The New York Times*, 22 Aralık, 2009). Kaliforniya, San Diego Üniversitesi araştırmacıları tarafından H.M.'nin beyninden 2000'den fazla kesit saklanmış ve dünyada çeşitli araştırmacıların da incelemesi için dijital olarak çoğaltılmıştır. Beyni çok ince dilimlere ayırma tekniği ile 21'inci yüzyılın bilgisayar teknolojisi birleştiğinde, beynin mimarisini daha önce hiç mümkün olmadığı kadar ayrıntılı ortaya çıkarma potansiyeli doğmaktadır. Teşekkürler H.M.

H.M.'in vasisi ve sinirbilimci Dr. Suzanne Corkin, biyografisini 2013'te yayımladı: *Permenent Present Tense: The Unforgetable Life of the Amnesic Patient H.M. (New York: Basic Books)*

Temel Kavram

Nomotetik yaklaşım çok sayıda katılımcıyı içerir ve grubun "ortalama" veya başka tipik değerlerini belirlemeye çalışır. Bu ortalama, gruptaki herhangi bir bireyin performansını temsil edebilir de etmeyebilir de. Burada araştırmacı daha ziyade bu grup ortalamasını kullanarak organizma performanslarını "genel olarak" yordamayı hedefler.

Başta Allport (1961) olmak üzere, birçok psikolog nomotetik yaklaşımın uygun olmadığını, bireyin çeşitli boyutlarda ortalama değerlerden oluşan bir derleme ile temsil edilemeyeceğini savunmuştur. Allport, bireyin içsel olarak tutarlı ilkeler doğrultusunda işlediğini, bu anlamda biricik ve kurallı olduğunu savunur. Dahası, bireyin daha ileri incelenmesi anlamına gelen **idiografik** yaklaşımın, psikolojik araştırmaların önemli bir amacı olduğunu ileri sürmüştür (bk. Smith ve ark., 1995).

Temel Kavram

Allport, idiografik yaklaşıma olan gereksinimi, klinik psikoloğun karşısındaki görevi betimleyerek anlatmıştır. Klinisyenin amacı "bütünü tahmin etmek değil, o bir kişinin ne yapacağını yordamaktır. Bu ideale ulaşırken, hesaba dayalı tahminler bazen yararlı olabilir; evrensel ve grup normları faydalıdır, ama sonuna kadar götürmezler." (s. 21). Allport, insan doğasını anlamak için kullanacağımız yaklaşımın sadece idiografik ya da sadece nomotetik olmaması gerektiğini, ikisi arasında bir dengeyi temsil etmesi gerektiğini öne sürmüştür. Vaka çalışması yönteminde görüldüğü gibi idiografik yaklaşım, "grup" yaklaşımının kaçırabileceği çeşitli küçük ayrıntı ve inceliği ortaya çıkarma gücüne sahip detaylı gözlem yapmaya olanak verir. Dahası, daha önce gördüğünüz gibi, vaka çalışması tipik olmayan bireyleri dikkatle inceleme yoluyla tipik ve ortalama davranış hakkında da bilgi verme yeteneğine sahiptir.

İLERİ ALIŞTIRMA

Bu alıştırmada, kısa betimlemenin ardındaki soruları cevaplamanız istenmektedir.

Bir arkadaşınızın bu yarıyılda Psikolojiye Giriş dersini aldığını ve size derste olanlara ilişkin tepkilerini anlattığını farz edelim. Günün konusu yetişkin gelişimiydi ve profesör, evlilik ve boşanma ile ilgili iki araştırma anlattı. Profesör, alanda uzman olan kişilerin yürüttüğü çalışmaların her ikisinin de mükemmel olduğunu vurguladı. Birinci çalışma iyi tanımlanmış bir evrenden seçilmiş geniş bir örneklemi içermekteydi. Bu çalışmanın bulguları, evliliklerin yarısından fazlasının boşanma ile sonlandığını, eşler arası sürekli anlaşmazlık ve ailede boşanma geçmişi olmasının, boşanmanın güvenilir yordayıcıları olduğunu göstermiştir. İkinci çalışma ise, bir çiftin evlilik ve aile danışmanı ile terapi deneyimlerinin uzun bir betimlemesiydi. Vaka çalışması, çiftlerin terapiye girmeden boşanmayı ciddi olarak düşündüklerini, bir yıllık terapi sonrasında ise evli kalmaya karar verdiklerini anlatmaktaydı. Profesör, evlilikte anlaşmazlık ve ailede boşanma geçmişinin onları boşanma riskine soktuğunu anlamaları ve bununla baş edebilmeleri için, çiftle çalışırken terapistin kullandığı birkaç özel tekniği açıkladı.

Profesör bu iki çalışmanın bulgularının nasıl birbirleriyle ilişkili olduğunu ve boşanma ile ilgili bunlardan nasıl bir sonuç çıkarılabileceğini tartışamadan dersin süresi bitti. Arkadaşınızın ders sonrasındaki kaygı ve sorularına nasıl cevap verirdiniz?

1 Arkadaşınızın sorularından biri, hangi çalışma bulgularına inanması gerektiğine ilişkindir. Birinci çalışma evlilikte anlaşmazlık ve aile geçmişinin boşanmaya yol açtığını söylemekteyken, ikinci çalışma bu faktörlerin boşanmayı gerektirmediğini gösteriyor. Arkadaşınız ikinci çalışmaya inanmaya eğilimli olduğunu belirtiyor. İkinci çalışmadaki kişisel örnekleri, birinci çalışmadaki rakamlarla desteklenen bulgulardan daha ikna edici bulduğunu söylüyor. Bu konuda siz ne düşünürsünüz?
2 Arkadaşınız aynı zamanda, bu çalışmaların, kendi hayatına ilişkin bir şeyler söyleyip söylemediği konusunu da sorgulamaktadır. Yani, bu çalışmaların bulgularına dayanarak, bir gün evlenmeyi seçtiğinde, kendisinin de boşanma deneyimi yaşayıp yaşamayacağını tahmin edebilir mi? Bu konuda siz ne düşünürsünüz?

Vaka Çalışması Yönteminin Dezavantajları

- Vaka çalışması yönteminde dış değişkenler kontrol edilemediği ve birden fazla "işlem" ardı ardına uygulanabildiği için araştırmacılar geçerli nedensel çıkarımlar yapamazlar.
- Gözlemci yanlılığı ve veri toplamadaki yanlılıklar vaka çalışması bulgularının yanlış yorumlamasına neden olabilir
- Vaka çalışması sonuçlarının genellenebilmesi, vakanın seçildiği evrenin değişkenliğine bağlıdır; bazı özellikler (örneğin, kişilik), diğerlerinden (örneğin, görsel keskinlik) daha fazla bireyler arası değişkenlik gösterir.

Neden-Sonuç Yargısına Varma Güçlüğü Buraya kadar bilimin bir amacının olguların nedenlerini keşfetmek (belirli bir olayın ortaya çıkışına ilişkin özel etmenlerin açıkça belirlenmesi) olduğunun farkına vardınız. Vaka çalışması yönteminin bir kısıtlılığı, bu çalışmalara bağlı olarak neden-sonuç yargılarının nadiren çıkarılabilmesidir. Bu kısıtlılık öncelikle araştırmacıların vaka

çalışmalarında dış değişkenleri kontrol edememesinden kaynaklanmaktadır. Dolayısıyla vaka çalışmalarında ortaya çıkan davranış değişimleri, birden fazla olası alternatif hipotezle açıklanabilir.

Örneğin, Suzan'ın öz yönetim eğitimi yöntemi ile tedavi edilmesini ele alın (bakınız, Kutu 9.2). Görünüşe bakılırsa Susan bu tedavi yönteminden faydalandı ama onun iyileşmesinin *nedeninin* bu tedavi yöntemi olduğundan emin olabilir miyiz? Birçok hastalık ve duygusal bozukluk, hiçbir tedavi uygulanmadan iyileşiyor. Vaka çalışması yapan araştırmacılar, bireylerin uygulamayı *almadan* iyileşmiş olabileceği alternatif hipotezini her zaman göz önünde bulundurmalıdır. Ayrıca, Susan'ın koşulları da birçok yönden onun iyileşmesini desteklemiş olabilir. Tedavisini, güven veren bir "klinik psikolog" üstlendi. Susan söz konusu tedavi nedeniyle değil, terapistinin içgörüsü ve kendi testlerinin sonuçlarından aldığı geri bildirim nedeniyle de kendisine karşı olan tutumunu değiştirmiş olabilir. Tedavinin bir parçası olarak terapist Susan'dan, onu endişeye sevk eden durumları sesli olarak veya zihninden kendi kendisine tekrarlamasını da istedi. Bu teknik, kendi başına etkili bir tedavi tekniği olan tekrara dayalı duyarsızlaşma tekniğine benzer (Rimm ve Masters, 1979).

Susan'ın iyileşmesinin kesin nedeninin söz konusu tedavi olduğunu, ikna edici şekilde gösteremeyiz, çünkü eşzamanlı olarak çeşitli tedaviler kullanıldı. Sonuç olarak, bu araştırmanın sonuçlarına dayanan çıkarımlar, daha ayrıntılı incelenineye kadar, geçici sonuçlar olarak kabul edilmelidir.

Yanlılığın Olası Kaynakları Bir vaka araştırmasının sonucu, çoğunlukla, hem katılımcı hem de gözlemci konumunda olan bir araştırmacının vardığı sonuçlara dayanır (Bolgar, 1965). Yani bir terapist, danışanının davranışını gözler *ve* tedavi sürecine katılır. Terapistin tedavinin başarılı olacağına olan inancının, onu motive ettiğini düşünebiliriz. Sonuç olarak terapist, iyi niyetli de olsa, danışanın davranışlarını yanlış gözleyebilir. Yanlı yorumun potansiyel kaynakları, vaka çalışması yöntemine özgü değildir. Daha önce gözlemci yanlılıklarını (Bölüm 4) ve deneyci yanlılıklarını (Bölüm 6) ele almıştık.

Bir vakanın bulguları, temelde gözlemcinin "izlenimlerine" dayanır (Hersen ve Barlow, 1976). Örneğin, terapist hastası Suzan'ın hedeflerine ulaşma becerisine ilişkin "duygularını" ve hayatı ile ilgili "kontrol hissini" nasıl anlattığını rapor etmiştir. "Davranışsal problem çözmeye dayalı bir değerlendirme yöntemini başarıyla kullandığını ve hedeflerine ulaşmak için stratejiler geliştirme konusunda oldukça ustalaştığını" bildirmiştir (s.304). Vaka analizi yönteminin ciddi bir zayıflığı, bulguların yorumlarının sadece gözlemcinin öznel izlenimlerine dayanmasıdır.

Vaka analizlerindeki yanlılık, bilginin sadece kişisel belgeler, seans notları ve psikolojik testler gibi kaynaklardan alındığında da ortaya çıkabilir. Dördüncü Bölümde bahsedildiği gibi, arşiv kayıtları birkaç yanlılık kaynağına açıktır. Dahası, bireyler kendileri hakkında bilgi verdiklerinde (öz bildirim), daha iyi görünmek için bilgiyi çarpıtabilir veya değiştirebilirler. Bu olasılık Susan'ın tedavisinde mevcuttu. Gelişim öz bildirimlerini abartıp abartmadığını bilmemizin bir yolu yoktur. Bir diğer potansiyel yanlılık kaynağı da raporların bireyin belleğine dayanmasıdır. Bilişsel psikologlar, özellikle uzun zaman önce

olmuş olayların hatırlanmasında belleğin doğru olmayabileceğini defalarca göstermişlerdir.

Tek Bir Bireyden Yola Çıkarak Genelleme Yapma Problemi Vaka analizlerinin en ciddi kısıtlılıklarından biri, vaka çalışması bulgularının genellenmesine ilişkindir. Bir bireyden elde edilen bulguları, daha büyük bir evrene ne ölçüde genelleyebiliriz? İlk karşılığımız, bir bireyin bulgusunun hiç genellenemeyeceği yönünde olabilir. Ancak tek bir vakadan genelleme becerimiz, vakanın seçildiği evrenin o konudaki değişkenliğine bağlıdır. Örneğin görsel algı çalışan psikologlar sıklıkla tek bir bireye dayanan bulgularını genelleyebilmektedirler. Görme ile ilgili araştırmacılar tüm insanlarda görme sistemlerinin aynı olduğunu varsaymaktadırlar. Bu nedenle, görsel sistemin nasıl çalıştığını anlamak için sadece bir veya birkaç vaka kullanılabilir. Buna karşılık, öğrenme, bellek, duygular, kişilik ve zihinsel sağlık gibi bazı psikolojik süreçler bireyler arasında daha fazla değişkenlik gösterirler. Evren içerisinde çok değişkenlik gösteren süreçleri incelediğimizde, bir bireyde gözlenenlerin tüm bireyler için geçerli olduğunu iddia etmek imkansızdır.

Dolaysıyla, terapistin öz yönetim tedavisi tekniğinin etkililiği ile ilgili sonuçlarını kabul etsek bile, bu tedavinin, çeşitli yönlerden Susan'a benzemeyen (örneğin, yaş, zekâ, aile geçmişi ve cinsiyet) diğer bireylerde de başarılı olup olmayacağını bilemeyiz. Grup yöntemlerinde olduğu gibi, önemli olan bir sonraki adım, çalışmayı çeşitli bireylerde tekrar (replike) etmektir.

Vaka Çalışmasına Dayanan Tanıklıklar Üzerine Eleştirel Düşünme

- Belirli bir uygulamanın etkililiği ile ilgili bireylerin kanıtları değerlendirildiğinde, vaka çalışması yönteminin sınırlılıklarına duyarlı olmak yararlıdır.

Vaka çalışmaları bazen yeni bulgulara ilişkin dramatik sonuçlar veya belirli bir tedavi başarısına ilişkin kanıtlar sunabilir. Medyada gördüğünüz ürün reklamlarını (örneğin, bilgi içeren reklamlar) düşünün. Eskiden aşırı kilolu olup X ürününü kullanarak çok kilo vermiş olan bir birey reklamda gösterildiğinde, buna kilo kaygısı yaşayan kaç kişi direnç gösterebilir? Vaka çalışmalarından sunulan kanıtlar çok ikna edici olabilirler. Bu, bilim camiası için hem avantaj hem de kısıtlılık oluşturur. Vaka çalışmaları yeni ve olağandışı bulgular göstererek bilim insanlarının kuramlarını yeniden gözden geçirmelerine neden olabilir veya yeni ve verimli araştırma alanlarına doğru yönlendirebilir. Bu açıdan vaka çalışmaları bilimin ilerlemesine yardım eder.

Vaka çalışmalarının kısıtlılığı ise, bulgularının eleştirilmeden kabul edilmesidir. Kilo vermeye veya hastalıktan iyileşmeye istekli bir birey vaka çalışmalarının sınırlılıklarını göz önüne almayabilir. Bunun yerine kanıtlar, bir çare için ışık ve umut olur. Az seçeneği olan (veya öyle düşünen) kişiler için bu nafile uğraş çok da mantıksız olmayabilir. Çoğunlukla insanlar, belirli bir işlemin neden kendileri için uygun olmadığını düşünmezler veya düşünmek istemezler.

TEK DENEKLİ (KÜÇÜK-N) DENEYSEL DESENLER

- Uygulamalı davranış analizinde, davranışın deneysel olarak analizi çerçevesinde geliştirilmiş yöntemler, sosyal açıdan uygun problemlere uygulanmaktadır.

Bu bölümün geri kalan kısmında tek denekli deneysel (küçük-*n*) desenlerini anlatacağız. Bu deneysel desenlerin kökleri, 1930 yıllarında B. F. Skinner tarafından geliştirilmiş olan davranışı inceleme yaklaşımında bulunmaktadır. Bu yaklaşım, *davranışın deneysel analizi* olarak adlandırılır.

Bu yaklaşımda, önceki bölümlerde tartışılan grup yaklaşımının aksine, örneklem genellikle tek bir denek veya az sayıda denekten (küçük-*n*) oluşmaktadır. Deneysel kontrol, koşullar düzenlenerek gösterilir; öyle ki bir bağımsız değişkenin değişimlenmesiyle, bireyin davranışı sistemli şekilde değişir. Skinner'ın (1966) yorumu şöyledir:

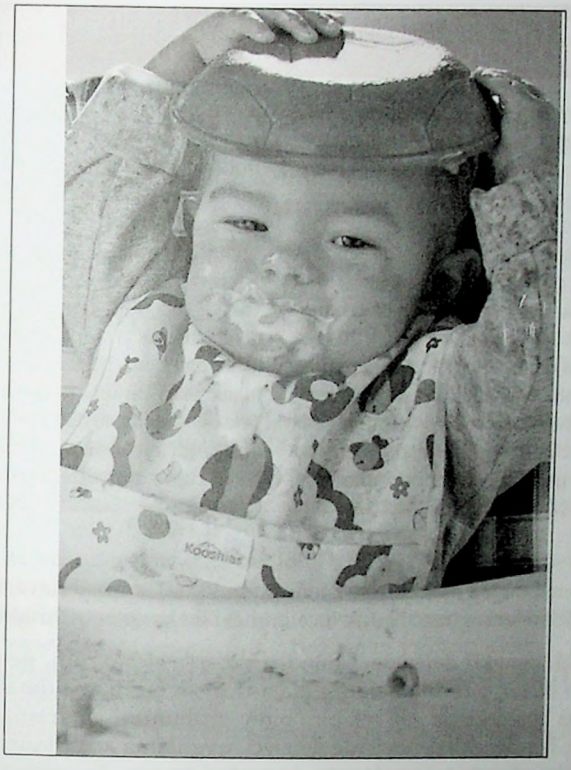
> Bin sıçanı birer saat incelemek veya yüz sıçanı onar saat incelemek yerine araştırmacı bir sıçanı bin saat incelemektedir. Bu süreç sadece bireyselliği tanıyan bir girişime uygun olmakla kalmaz, aynı zamanda araçların ve araştırmacının zaman ve enerjisini kullanmada da aynı derecede etkindir. İstikrar ve tekrarlanabilirlik, kullanılan yöntemlerle değil ulaşılan kontrolle en iyi şekilde ölçülür ki bu, davranışın deneysel analizinin çok kolay geçebildiği bir sınavdır (s. 21).



Tek denekli deneysel desende çoğunlukla, yapılması gereken çok az istatistiksel analiz vardır. Deneysel bir değişkenin (tedavi) etkilerine dair sonuçlara, tipik olarak, şöyle varılır: Deneysel tedavinin uygulanması ve sona erdirilmesine bağlı olarak davranışın da değişip değişmediğini gözlemek için, davranış kayıtlarına bakılır. Bu nedenle, davranışın tanımlanması, gözlenmesi, ve kaydedilmesi hususlarına çok önem verilmesi beklenir. Davranış, güvenilir şekilde gözlenip kaydedilecek şekilde açık ve nesnel olarak tanımlanabilir mi? Davranışın sürekli (birikimli) bir kaydı mı tutlacak, yoksa gözlemler düzenli aralıklarla mı yapılacak? Her ne kadar tepkinin sıklığı davranışın yaygın bir ölçüsü ise de, davranışın süresi veya diğer özellikleri de bazen ölçülür. Üstelik, bu bölümde daha sonra göreceğiniz gibi, bazen istatistiksel sorunlar ortaya çıkar; davranışın kaydındaki aşırı değişkenlik gibi; bunun üstesinden gelmek gerekir. Tek denekli araştırma deseniyle ilgili diğer istatistiksel meseleler, bizim kısa girişimizin çok ötesine geçer (bk. örneğin; Kratochwill ve Levin, 1992; Parker ve Brossart, 2003).

Uygulamalı davranış analizinde, davranışın deneysel analizi yöntemleri, ilgili sosyal problemlere uygulanmaktadır. Bu uygulamalar sıklıkla *davranış değiştirme* olarak adlandırılır, ancak klinik popülasyonlara uygulandığında *davranış terapisi* terimi tercih edilmektedir (Wilson, 1978). Davranış terapisi, birçok psikolog tarafından klinik tedavi için psikodinamik modele dayalı terapiden daha etkili bir yaklaşım olarak görülmektedir. Davranış terapisi,

ŞEKİL 9.2 Uygulamalı davranış analizi, çocuk ve yetişkinlerin uyumsuz davranışlarını kontrol etme yöntemlerini incelemek için kullanılır



problemlerin bilinçaltı kökenlerine inmek yerine gözlenen davranışa odaklanmaktadır. Davranış değiştirmenin ve davranış terapisinin kekemelerin, normal ve zihinsel engelli çocuk ve yetişkinlerin, psikiyatrik hastaların ve daha bir çoklarının davranışını nasıl değiştirdiğine dair sayısız çalışma yayımlanmıştır (bk. Şekil 9.2). Uygulamalı davranış analizi yaklaşımları, okul psikologları tarafından da eğitim ortamlarında başarıyla kullanılmıştır (bk. Kratochwill ve Martens, 1994). *Journal of Applied Behavioral Analysis* adlı bilimsel dergi, bu tür yayımlanmış çalışmalar için temel bir kaynaktır.

Tek Denekli Deneylerin Özellikleri

- Araştırmacılar tek denekli deneylerde bir bağımsız değişkeni değişimler; bu nedenle bu desenler vaka çalışmalarından daha titiz kontrole izin verir.
- Tek denekli deneylerde, işlem öncesi bireyin davranışının nasıl olduğunu betimleyebilmek (ve gelecekte nasıl olacağını tahmin edebilmek) amacıyla önce temel düzey gözlemler kaydedilir.

KUTU 9.4

TEK DENEKLİ DESENLERİN GRUP DESENLERİNDEN ÜSTÜN YÖNLERİ: AZ FAZLA OLABİLİR

Bir grup deneyinin bulgusu, davranışı şekillendirmede hangi tedavilerin "genel olarak" etkili olduğuna ilişkin önerilere öncülük edebilir. Ancak grup ortalamasına dayanarak, tedavinin belirli bir birey üzerinde etkisinin ne olacağını söylemek mümkün değildir. Tek denekli desenler, grup yöntemlerinin bulgularının belirli bir hasta veya bireye uygun olup olmadığını sınamak için uygundur (Kazdin, 2011).

Tek denekli deneylerin çok gruplu deneylerden bir başka üstünlüğü de, klinik araştırmalarda ortaya çıkan tedaviyi vermeme etik sorununu (withholding treatment) içermektedir. Çoklu grup desenlerinde, potansiyel olarak yararlı bir tedavi, iç geçerlik gereksinimlerini karşılamak için oluşturulan kontrol grubuna verilmemektedir. Tek denekli deneysel desenlerde ise "işlem var" ve "işlem yok" koşulları aynı denek üzerinde değişimlendiği için, tedavinin verilmemesi problemi ortadan kalkar. Dahası, klinik araştırmacılar çoklu grup deneyleri için yeterli sayıda hastaya ulaşmada zorluk yaşarlar. Örneğin, bir klinik psikolog, kapalı alan fobisi (klostrofobi) tanısı alan birkaç hastaya ancak ulaşabilir. Tek denekli deneyler, sadece birkaç katılımcının mevcut olduğu durumlara ilişkin pratik bir çözüm sunmaktadırlar.

- Temel düzey davranış ve işlem (uygulama) sonrası davranış, kaydedilmiş gözlemlerin görsel olarak incelenmesi ile karşılaştırılmaktadır.

Temel Kavram

Tek denekli deneyde, isminden de anlaşıldığı gibi, genellikle bir veya en fazla birkaç bireyin davranışındaki değişime odaklanılmaktadır. Ancak daha sonra göreceğimiz gibi tek bir "birey grubunun" davranışına da odaklanılabilir. Tek-denekli deneylerde deneyci, uygulama koşullarını sürekli izlenen bir tek birey üzerinde karşılaştırır. Diğer bir deyişle, ilgilenilen bağımsız değişken tek bir birey için sistematik olarak değişimlenir. Tek-denekli deneysel desenler göreceli olarak daha kontrolsüz olan vaka çalışması yöntemine önemli bir alternatiftir (Kazdin, 2011). Kutu 9.4'te anlatıldığı gibi tek denekli deneylerin çok-gruplu deneylerden de daha avantajlı yönleri bulunmaktadır.

Temel Kavram

Tek denekli deneylerde ilk basamak genellikle gözlem veya **temel düzey aşamasıdır**. Bu aşama sırasında araştırmacılar hiçbir işlem yapılmadan deneğin davranışını kaydederler. Genellikle klinik araştırmacılar, bir gün veya bir saat gibi belli bir zaman birimi içinde, bir hedef davranışın sıklığını ölçerler. Örneğin araştırmacı, aşırı utangaç bir çocuğun 10 dakikalık bir görüşme sırasında kaç kez göz teması yaptığını, bir migren hastasının her hafta kaç kez migren ağrıları rapor ettiğini, veya kronik bir kekemenin bir dakikada kaç kez durakladığını kaydedebilir. Temel düzey yöntemini kullanarak araştırmacılar, işlem (tedavi) öncesindeki davranışı betimleyebilirler. En önemlisi temel düzey araştırmacılara, tedavinin olmaması durumunda davranışın ne olacağını *yordama* imkanı verir (Kazdin, 2011). Elbette davranış gerçekten izlenmediği sürece araştırmacılar gelecekte davranışın ne olacağını kesin olarak bilemezler,

ama temel düzey hiçbir işlemin yapılmadığı durumda davranışın ne olacağına ilişkin tahmin imkanı sağlayabilir.

Araştırmacılar bireyin davranışının göreceli olarak sabit olduğunu, yani kayıt aralıklarında küçük dalgalanmalar olduğunu gözledikleri zaman, bir işlemi uygulamaya koyarlar. Bir sonraki basamak, temel düzeyde kullanılan ölçütlerin aynısını kullanarak bireyin davranışlarını kaydetmektir. Araştırmacılar, uygulama sonrasında gözlenen davranışı temel düzey performansı ile karşılaştırarak tedavinin etkisini belirleyebilirler. Tedavinin etkisi davranışsal kaydın grafiği kullanılarak kolaylıkla görülebilir. Davranış deneysel uygulamadan sonra nasıl değişti? Tedaviden sonraki davranış ile tedavi olmadığı durumda tahmin edilen davranıştaki farkı görsel olarak karşılaştırarak işlemin davranışı etkili biçimde değiştirip değiştirmediğine ilişkin çıkarım yapabiliriz. En yaygın kullanılan Tek denekli deneysel desenler ABAB deseni ve çoklu temel düzey desenleridir.

Özel Deneysel Desenler

- ABAB deseninde temel düzey (A) ve uygulama (B) aşamaları işlemin davranış üzerindeki etkisini belirlemek için yer değiştirir.
- Davranışın, işlemin uygulanması ve geri çekilmesi ile sistematik olarak değişmesi durumunda araştırmacılar, işlemin davranışı değiştirdiğine karar verirler.
- ABAB deseninde işlem geri çekildiğinde davranış temel düzeye geri dönmezse, işlemin davranışa etkisini yorumlamak zor olur.
- Etik kaygılar psikologların ABAB deseni kullanmalarını engelleyebilir.
- Çoklu temel düzey desenlerinde, değişimin yalnızca işlem uygulanmasından sonra ve birden fazla temel düzeyde ortaya çıkması, işlemin etkililiğini gösterir.
- Çoklu temel düzeyler bireyler, davranışlar veya ortamlar üzerinden gözlenebilir.
- Çoklu temel düzey desenlerinde, deneysel işlem öncesinde temel düzeyde bir değişim görüldüğünde, işlemin nedensel etkisini yorumlamak zordur. Böyle bir durum işlem etkilerinin genellenmesi durumunda ortaya çıkabilir.

Temel Kavramlar

ABAB Deseni Araştırmacılar, davranışın sistematik olarak değiştiğini göstermek için "işlem var" ve " işlem yok" koşulları değişimlediklerinde, **ABAB desenini** kullanmaktadırlar. Başlangıçtaki temel düzey aşamasını (A), bir işlem aşaması takip eder (B); sonrasında temel düzeye dönüş (A) ve bir kez daha işlem aşaması (B) gelir. İşlem ikinci A aşamasında geri çekildiği için davranışta ortaya çıkan herhangi bir gelişmenin tersine dönmesi olasıdır; bu desen aynı zamanda **tersine desen** adıyla da anılır. ABAB desenini kullanan araştırmacı, işlem değişkeninin sunumundan hemen sonra davranışta değişim olup olmadığını (ilk B), işlem geri çekildiğinde davranışın tersine dönüp dönmediğini (ikinci A) ve işlem tekrar verildiğinde (ikinci B) davranışın gelişip gelişmediğini gözlemler. Davranış işlemin verilmesi ve geri çekilmesinden sonra değişiyorsa, araştırmacı işlemin davranıştaki değişime neden olduğuna ilişkin önemli bir kanıt sağlamış olur.

ABAB deseni, 8 yaşındaki ağır zihinsel engelli bir kız çocuğunun uyumsuz davranışını azaltmada yüz kapatmanın etkilerini değerlendirmek için kullanılmıştır (Horton, 1987). Yüz kapatma, istenmeyen davranış ortaya çıktığında yüzün kapatılmasını (örneğin, yumuşak bir kumaş parçası ile) içeren hafif itici (aversive) bir tekniktir. Önceki çalışmalar, yüze tokat atma gibi kendini yaralayıcı davranışların sıklığını azaltmada etkili bir teknik olduğunu göstermiştir. Bu çalışmada, küçük çocuğun yemek sırasında kaşık çarpma davranışını azaltıp azaltmayacağının belirlenmesi amaçlanmıştır. Kaşık çarpma, öğrencinin devam ettiği zihinsel engelliler okulunda sınıf arkadaşlarıyla birlikte yemek yemesini engellemekteydi. Çarpma, sadece gürültü yarattığı için değil, aynı zamanda yiyeceklerin etrafa saçılmasına neden olduğu ve çoğu zaman kaşığın yere düşmesine yol açtığı için bozucu bir davranıştı.

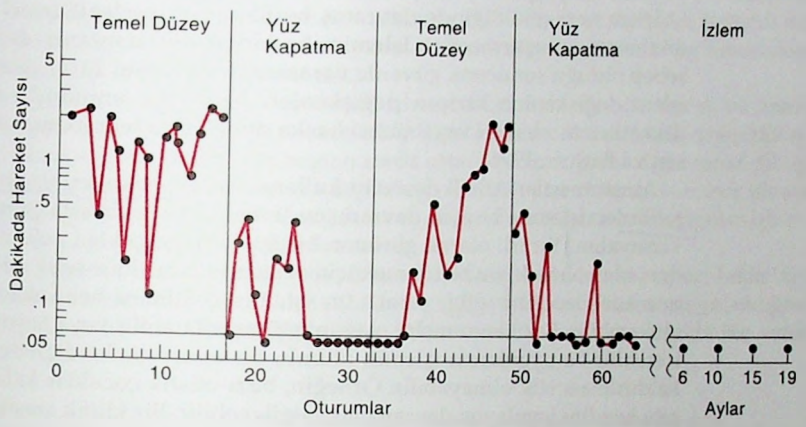
Normal kaşıklama davranışından ayırt etmek için kaşık çarpmanın açık bir tanımı yapıldı. Daha sonra bir sağlık görevlisi, gözlem yapmak ve işlemi uygulamak için eğitildi. Her 15 dakikalık yeme seanslarında kaşık çarpma sıklığı sayılarak büyüklük değerlendirildi. Başlangıç veya temel düzey aşamasında sağlık görevlisi sıklığı kaydetti ve tepkinin her ortaya çıkışında kızın bileğinden tutarak elini tabağına getirdi ve "vurmak yok" dedi. Uygulama videoya kaydedildi; güvenirliğin kontrolü amacıyla bağımsız bir gözlemci tarafından izlenerek davranış sıklığı kaydedildi. Gözlemciler arası geçerlik yaklaşık %96 olarak bulundu. Temel düzey aşaması 16 gün boyunca yürütüldü.

17. günde birinci işlem aşaması başlatıldı ve 16 gün sürdürüldü. Her kaşık çarpma gözlendiğinde sağlık görevlisi "vurmak yok" geri bildirimini vermiş ve kızın elini tabağa döndürmüştür. Ancak bu kez sağlık görevlisi ek olarak kızın tüm yüzünü 5 saniye boyunca havlu kumaş ile kapatmıştır. Yüz kapatmanın sonlandırılması ise kaşık çarpma davranışının 5 saniye yapılmaması koşuluna bağlanmıştır. Birinci işlem aşamasını, ikinci bir temel düzey aşaması ve sonra yeni bir işlem aşaması takip etmiştir. İşlem sonrası gözlemler ise 6, 10, 15 ve 19'uncu aylarda yapılmıştır.

Şekil 9.3 kızın kaşık çarpma davranış sıklığını, değişen temel düzey ve işlem aşamalarının bir fonksiyonu olarak göstermektedir. Yüz kapatma, yalnızca işlem aşamalarında davranışı azaltmakta etkili olmamış, izleme gözlemleri aylar sonra kaşık çarpmanın görülmediğini de ortaya koymuştur. Son işlem aşamasının ardından kızın evde veya okulda yemek saatinde doğrudan gözetime ihtiyacı kalmamış ve arkadaşlarıyla birlikte yemesine izin verilmiştir. Yüz kapatma işlemi, uygulanan tek işlem olmuş ve şekil 9.3'ün görsel olarak incelenmesi davranışın, işlemin yapılması ve çekilmesi ile birlikte değiştiğini göstermiştir. Daha az girişimsel başka işlemlerin başarısız olduğu genç bir çocuğun uyumsuz davranışını kontrol etme durumunda, yüz kapatma tekniği başarılı bir işlem olmuştur.

ABAB Desenleriyle İlişkili Yöntemsel Sorunlar ABAB işlemi kapsamında bazen ortaya çıkan temel bir yöntemsel sorun da Şekil 9.3'e tekrar bakılarak gösterilebilir. Yüz kapatma uygulamasının geri çekildiği ikinci temel düzey basamağında,

ŞEKİL 9.3 Çalışmanın temel düzey, işlem ve izlem evrelerinde kaşık çarpma sıklığı [Horton'dan (1987) uyarlanmıştır]



kaşık çarpma artmıştır. Yani işlem aşamasında gözlenen gelişme tersine dönmüştür. İşlem ortadan kalktığı halde kaşık çarpma davranışı düşük kalmaya devam etseydi ne olurdu? İkinci temel düzeyde davranış, başlangıçtaki temel düzey durumuna geri dönmez ise, araştırmacı işlemin etkililiği konusunda nasıl bir sonuca varabilir?

ABAB deseninde davranışın temel düzeye geri dönmemesinin birkaç nedeni vardır (Kazdin, 2011). Birinci neden, işlemin bireylere yeni beceriler öğrettiği durumlarda, yani işlem gelişmeye yol açtığında, mantıklı olarak davranışın değişmesi beklenmez. Örneğin, bir araştırmacı engelli bir bireye iş yerine nasıl ulaşacağını öğretebilir. Beceri öğrenildikten sonra onu unutması (temel düzeye geri dönmesi) beklenmez. Bu sorunun çözümü açıktır: Araştırmacılar, işlem geri çekildiğinde hedef davranışların temel düzeye dönmesinin beklenmediği durumda ABAB desenini kullanmamalıdır.

Davranışın temel düzeye dönmemesinin bir başka olası nedeni, işlem dışındaki *bir başka* değişkenin davranıştaki değişime yol açmış olması olabilir. Örneğin, birey işlem sırasında personelin veya başkalarının ilgi artışına maruz kalabilir. İkinci temel düzey aşamasında belirli işlem geri çekildiği halde dikkat sürdürülürse, davranıştaki gelişmenin sürmesi de olasıdır. Bu açıklama, işlem değişkeni ile bir başka değişkenin (örneğin, ilgi) arasında karıştırıcı durumu öne sürmektedir.

İşlemin başlangıçta davranışı geliştirmesi, ama daha sonra yeni davranışın kontrolünü diğer değişkenlerin üstlenmesi de olası bir durumdur. Bir kez daha dikkat kritik bir değişken olabilir. Aile ve arkadaşlar davranışın değiştiğine tanık olduklarında, bireye daha fazla ilgi gösterebilirler. İnsanların kilo kaybettikleri veya sigarayı bıraktıkları zaman karşılaştıkları övgüyü düşünün. Bu olumlu pekiştirme (ilgi) işlemin başlattığı davranış değişimini devam ettirebilir.

İşlem geri çekildiğinde ilgi devam ederse, davranışın temel düzeye dönmesini beklemeyiz.

İşlem geri çekildiğinde davranış herhangi bir nedenle temel düzeye dönmezse, araştırmacılar işlemin başlangıçtaki davranış değişimine sebep olduğu sonucuna güvenle varamazlar (Kazdin, 2011). Araştırmacı işlem değişkenine karışan değişkenleri belirleme umuduyla durumu dikkatlice incelemeli veya süreci başka deneklerle tekrarlamalıdır (Hersen ve Barlow, 1976).

Araştırmacılar ABAB desenini kullanırken bir etik sorunla da karşılaşabilirler. İşlemin, bireyin davranışını temel düzeye kıyasla geliştirdiğini varsayalım. Yararlı olarak görünen bir işlemin, gerçekten gelişmenin nedeni olup olmadığını belirlemek için ortadan kaldırılması etik midir? Tahmin edebileceğiniz gibi, yararlı bir işlemin çekilmesi her vakada meşru olmayabilir. Bazı davranışlar olağanüstü zayıflatabilir veya hayati tehlike içerebilir ve uygulanan işlemin olumlu bir etkisi gözlendiğinde ortadan kaldırılması etik olmayabilir. Örneğin, bazı otistik çocuklar kafa çarpma gibi kendini yaralayıcı davranışlar sergileyebilir. Bir klinik araştırmacı bu davranışların sıklığını azaltmayı başarırsa, ABAB desenin gerekliliklerini karşılamak için işlemi geri çekmesi etik dışıdır. Neyse ki, işlemin geri çekilmesini içermeyen ve bu durumlarda uygulanabilecek bir tek denekli deneysel desen mevcuttur: Çoklu temel düzey deseni.

Çoklu Temel Düzey Deseni Çoklu temel düzey deseni de temel düzey ve işlem basamaklarından yararlanır, ancak ABAB desenindeki geri çekme tekniğini kullanmaz. Adından da anlaşılacağı gibi, araştırmacılar çoklu temel düzey deseni kullanırken birden fazla temel düzey saptarlar. Çoklu temel düzey deseni, işlemin uygulanmasından sonra davranışın birden fazla temel düzeyde değiştiğini göstererek işlemin etkililiğini ortaya koyar.

Temel Kavram

Çoklu temel düzey desenini *Journal of Applied Behavioral Analysis* adlı dergiden bir örnek vererek gösterelim. Araştırmacılar seçici mutizm (suskunluk) gösteren 9 yaşındaki bir kızın tedavisi için **ortamlar arası çoklu temel düzey deseni** kullanmışlardır (Lang, Regester, Mulloy, Rispoli ve Botout, 2011). Seçici mutizm, çocuğun bazı ortamlarda konuştuğu, diğerlerinde ise sessiz kaldığı bir gelişimsel bozukluktur. Örneğin, çocuk evde konuşabilir ancak okulda veya diğer topluluk ortamlarında (örneğin; restoran, park) konuşmayabilir. Leslie evde sıklıkla konuşuyor, ancak ev dışı birçok ortamda konuşmuyordu. Mutizmi bir yıldan uzun sürmüş ve yaz kampı, yüzme dersleri gibi birçok sosyal ortamda zorluklara yol açmıştı.

Leslie'nin evinde yürütülen tedavi rol yapma, video ile kendini model alma ve pekiştirmeyi içermekteydi. Tedavi seanslarında deneyci Leslie'nin soruları cevaplayarak çeşitli ortamlarda rol yapmasını istemekteydi (örneğin, restoran: "Ne ısmarlamak istersin?"). Daha sonra kendini videoda izlemiş, yanıtları ve konuşmayı başlattığı için övgü almıştır (ayrıntılar için bk. Lang ve ark., 2011).

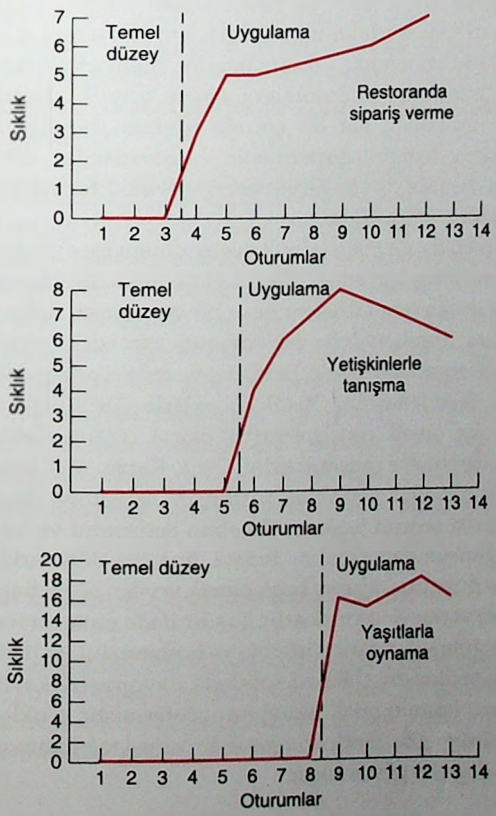
Tüm çoklu-temel düzey desenlerinde olduğu gibi bu desende de iki veya daha fazla temel düzey saptanır ve daha sonra işlem farklı zamanlarda uygulanır. Bu çalışmada temel düzey üç farklı ortamda gözlenmiştir: Bir restoranda, yeni yetişkinlerle tanıştığı bir ofis binasında ve yeni yaşıtlarıyla

oyun masasında. Davranış tüm ortamlarda izlenmiş ve işlem bu ortamlarda teker teker uygulanmıştır. Deneyci işlem öncesi (temel düzey) ve sonrasında annesiyle beraber Leslie'ye her ortamda eşlik etmiş ve üç farklı bağımlı değişkeni kaydetmiştir: Sorulara yanıtlar, konuşmayı başlatmalar ve tutukluklar (örneğin, iletişimde başarısızlık).

Çoklu temel düzey deseninde işlem *bir* ortamda sunulur ancak bu arada diğer ortamlarda izleme devam eder. İşlem etkili ise, davranıştaki değişiklik işlemin uygulanmasından hemen sonra ortaya çıkacak, ancak diğer temel düzey ortamlarında gözlenmeyecektir. Dolayısıyla, *çoklu temel düzey deseninde işlemin etkililiğine ilişkin anahtar kanıt, davranış değişiminin sadece işlemin uygulandığı ortamda ortaya çıktığının gösterilmesidir.*

Şekil 9.4'te bağımlı değişkenlerden birisinin verisi sunulmuştur: Leslie'nin sorulara yanıtları. Grafiklerdeki dikey çizgiler, işlemin ne zaman uygulandığını göstermektedir. Görüldüğü üzere, Leslie temel düzey seanslarında hiç yanıt

ŞEKİL 9.4 Leslie'nin üç sosyal ortamdaki tepki sıklığını gösteren çoklu temel düzeyler [Lang ve arkadaşlarından (2011) uyarlanmıştır, Şekil 1]



vermemiştir. Her ortamda işlem uygulanmaya başlandıktan sonra yanıt sıklığı artmıştır. Bulgular, bu tedavinin Lesli'nin mutizmini üç ortamda da başarılı bir biçimde geliştirdiğini göstermektedir.

Çoklu-temel düzey deseninin temel düzeylerin farklı bireylere, aynı bireyde farklı davranışlara veya aynı bireyde farklı ortamlara uygulandığı birkaç çeşidi bulunmaktadır. Karmaşık görünmesine rağmen çoklu-temel düzey desenleri sıklıkla kullanılmakta ve kolaylıkla anlaşılmaktadır.

Temel Kavram

Bireyler arası çoklu-temel düzey deseninde öncelikle farklı bireylerde temel düzeyler tespit edilir. Bir bireydeki davranış sabitlendiğinde işlem sırayla diğer bireylere uygulanır. *Journal of Child Psychology and Psychiatry* dergisinde yayımlanmış olan bir çalışmada bu desen kullanılmıştır (Whalen ve Schreibman, 2003). Araştırmacılar birkaç otistik çocukta "birleşik dikkat" davranışını değiştirmeyi amaçlamışlardır. Birleşik dikkat, sosyal bir ortamda hem nesne hem de kişiye dikkat edebilme becerisine verilen addır ve araştırmalar otizmin birleşik dikkatte bozulmayla ilişkili olduğunu göstermektedir. Tedavi, çocuklara birleşik dikkat davranışını öğreten davranış şekillendirme yöntemlerini içermektedir. Şekil 9.4'e baktığınızda, her temel düzeyin farklı bir ortamı değil, faklı bir bireyi temsil ettiğini düşünün. Tüm çoklu-temel düzey desenlerinde olduğu gibi, işlem her temel düzeye (burada her bir bireye) farklı zamanda sunulmaktadır. İşlem etkili ise her bir bireyde işlemin sunumunun hemen ardından değişim ortaya çıkacaktır. Çoklu-temel düzey deseni bulguları, her bir çocuğa işlemin uygulanmasından sonra birleşik dikkatin geliştiğini göstermiştir. Araştırmacılar, dikkat becerilerinin tedavi ortamı dışında da sürdürülmesi için tedavi tekniklerinin ebeveynlere de öğretilmesini önermişlerdir.

Temel Kavram

Üçüncü bir tür çoklu-temel düzey deseni ise aynı bireyde iki veya daha fazla farklı davranış için temel düzey belirlemektedir. **Davranışlar arası çoklu temel düzey deseninde** bir işlem önce bir davranışa, daha sonra sırayla diğer davranışlara uygulanır. Her bir davranış için işlem uygulandıktan hemen sonra değişim gerçekleştiğinde, davranış ve tedavi arasındaki nedensel ilişkiye kanıt olarak değerlendirilir. Otistik çocuklarla yürütülen bir başka çalışmada araştırmacılar, otistik gençlere sosyal olarak uygun afektif (duygu içeren) davranışlar öğretmeye çalışmışlardır (Gena, Kantz, McClannahan ve Poulson, 1996). Otistik çocuklar genellikle uygunsuz duygusal davranışlar göstermekte ve bu durum onların başkalarıyla etkin iletişimini ve kişiler arası ilişkiler kurma olanağını sınırlamaktadır. Tedavi, üç veya dört farklı davranış kategorisinde uygun duygusal tepkiye bağlı olarak verilen sözel övgü ve fiş (ödülle değiştirilebilen) vermeyi içermektedir. Takdir ifade etme, favorilerden konuşma, garipliklere gülme, sempati gösterme ve hoşlanmamayı ifade etme hedef davranışlar olarak seçilmiştir. Çalışma sonunda davranışsal kayıtların görsel olarak incelenmesi, işlemin etkili olduğunu göstermiştir. Çoklu temel düzey deseninin gerektirdiği gibi, farklı duygusal davranışlar işlemin uygulanmasından hemen sonra değişim göstermiştir.

Çoklu Temel Düzey Desenine İlişkin Yöntemsel Sorunlar

- Kaç tane temel düzey gereklidir?

Tek denekli araştırmanın birçok alanında olduğu gibi "kaç tane temel düzeye ihtiyacım var?" sorusuna cevap vermeye yarayan belirli ve sağlam kurallar yoktur. En düşüğü iki temel düzeydir ama çoklu temel düzey deseninde üç veya dört temel düzey önerilmektedir.

- Davranış müdahale öncesinde değişirse ne olur?

Her tip çoklu temel düzey deseninde, işlem uygulanmadan önce temel düzey esnasında davranışta bir değişim meydana gelirse bu bir sorun oluşturur. Temel düzeyde görülen bu olgunlaşmamış değişimlerin nedeni her zaman açık değildir. Çoklu temel düzey deseninin mantığı, işlem sonrasında davranışın değişmesidir. Bu nedenle müdahale öncesi temel düzeyde bir değişim meydana geldiğinde, işlemin etkili olduğu sonucuna varmak zordur. Eğer işlem öncesi değişiklikler sadece bir temel düzeyde ortaya çıkmışsa (özellikle değişimin durumsal veya işlemsel faktörlere bağlı olduğuna ilişkin bir açıklama varsa), çoklu temel düzey deseninin bir miktar güvenle yorumlanmasına devam edilebilir. Örneğin, bireyler arası çoklu temel düzey deseninde işlemi çalışmanın daha geç bir evresinde alan kişiler, diğer bireylerin uyum gösterdiklerini görebilir ve onların davranışlarını taklit edebilirler. Bu durumda işlemi en sonda alan bireylerde işlem öncesi görülen davranış değişiklikleri açıklanabilir.

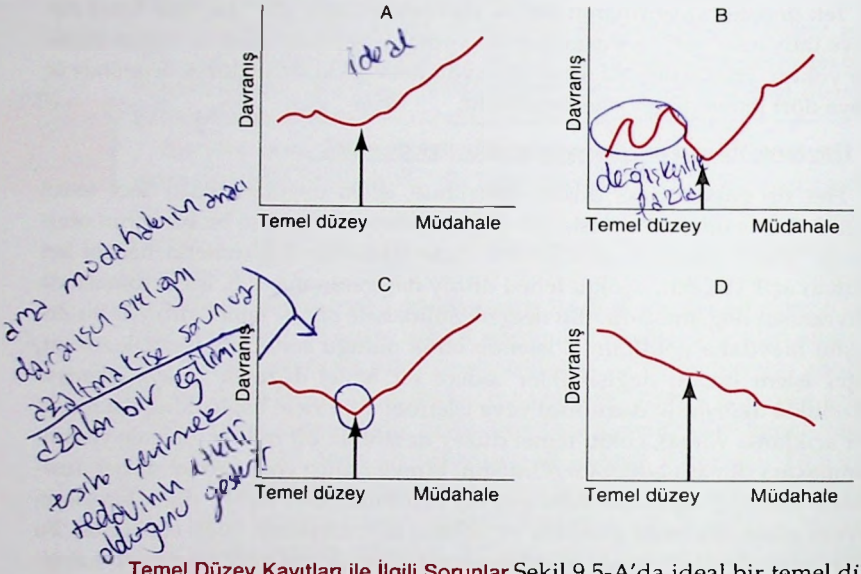
- İşlem diğer davranış ve ortamlara genellenirse ne olur?

Çoklu temel düzey deseninde ortaya bazen çıkabilen bir sorun da bir davranıştaki değişimin diğer davranışlara veya ortamlara genellenmesidir. Genelleme ile ilgili problemlerle uğraşırken araştırmacılar, "Bir gram koruyucu tedbir, bir kilo tedaviye eşittir." deyişini akıllarında tutmalıdır. Bir bireyin davranışını değiştirmek diğerlerinin davranışını etkiliyorsa, bir ortamdaki davranış diğer ortamdaki davranışı etkiliyorsa veya bir tip davranışın değiştirilmesi diğer davranışları etkiliyorsa, çoklu temel düzey deseni tekrar düzenlenmeli veya belki de terk edilmelidir (Kazdin, 2011). Açık olan şudur ki, çoklu temel düzey deseninde işlemin etkili olduğu sonucuna varabilmek için her temel düzeyde işlemin uygulanmasından sonra değişimin ortaya çıkması gereklidir.

Tüm Tek Denekli Desenlerin Ortak Sorunları ve Sınırlılıkları

- Temel düzey aşamasında, davranışta aşırı değişkenlik veya artan, azalan eğilimler olduğunda, işlemin etkisini yorumlamak zor olabilir.
- Tek denekli desenlerde dış geçerliğin düşük olması problemi, küçük grupların sınanması ile giderilebilir.

ŞEKİL 9.5 Bir davranış şekillendirme programının temel düzey ve müdahale aşamalarının olası ilişkilerini gösteren davranışsal kayıt örnekleri. Grafikteki ok, müdahalenin başlangıcını göstermektedir.



Temel Düzey Kayıtları ile İlgili Sorunlar Şekil 9.5-A'da ideal bir temel düzey kaydı ve müdahaleye ilişkin tepki gösterilmektedir. Temel düzeyde davranış çok durağanken, müdahalenin uygulanmasının hemen ardından davranış değişmiştir. ABAB'nin ilk basamağında veya çoklu temel düzey deseninde bulgular böyle olduğu zaman, işlemin davranışı değiştirmede etkili olduğunu gösterme yolunda olurduk. Ancak B panelindeki temel düzey ve müdahale aşamalarını göz önüne aldığımızda, istenen davranış müdahale sonrasında artış gösteriyor olsa da, temel düzeyde değişkenlik çok yüksektir. Burada işlemin mi değişimi ortaya çıkardığı, yoksa davranışın tesadüfen mi yükseldiğini söylemek zordur. Genel olarak temel düzeyde değişkenliğin yüksek olduğu durumda işlemin etkililiğine karar vermek güçtür.

Temel düzeydeki değişkenliğin ölçüm hatası, ortamlardaki kontrol edilmeyen faktörler ve denekler arası farklılıkların da dahil olduğu birçok kaynağı olabilir (Kazdin, 2011). Aşırı değişkenlik ile başa çıkmanın bir yolu, değişkenlik üreten etmenlerin aranması ve ortadan kaldırılmasıdır. Örneğin bir personelin varlığı, bir psikiyatri hastasında davranış değişikliğine neden oluyor olabilir. Bir başka yaklaşım ise "beklemektir"; temel düzeyde istikrar sağlanana kadar ölçüm almaya devam etmektir. Elbette bunun olup olmayacağı ve ne zaman olacağını tahmin etmek mümkün değildir. Ancak müdahaleyi istikrar olmadan uygulamak da sonucun yorumunu tehlikeye sokmaktadır.

Şekil 9.5-C'de temel düzeyin artan veya azalan bir eğilim göstermesi sonucu ortaya çıkabilecek problem gösterilmektedir. Eğer müdahalenin hedefi davranışın sıklığını arttırmak ise, C panelinde görülen azalan eğilimin yorumlanması

sorun içermez. Azalan bir eğilimi tersine çeviren bir müdahale işlemin etkili olduğunun bir kanıtı olabilir. Ama müdahalenin amacı davranışın sıklığını azaltmak ise, burada daha ciddi bir problem var demektir. Bu durum D panelinde gösterilmektedir. Burada, temel düzey aşamasında azalan bir eğilim ve işlem aşamasında devam eden bir sıklık düşüşü görmekteyiz. İşlemin etkili olup olmadığını bilmek zordur çünkü, işlem aşamasında görülen azalma işleme bağlı olabileceği gibi temel düzeyde başlayan azalmanın devamı da olabilir. İşlemin beklenen etkisi temel düzeyde görülen eğilim ile aynı yönde olduğunda, işlemin etkili olduğu sonucuna varabilmek için, müdahaleyi takip eden değişimin D panelinde görülenden çok daha belirgin olması gerekir. Davranışsal kaydı görsel olarak incelerken "belirgin" değişikliğin ne olduğunu söylemek genellikle zordur (bakınız, Parsonson ve Baer, 1992). Bu koşullarda, hedef davranışa ilişkin gözlemlere ek olarak "normal" kişilerle karşılaştırma yapmak veya bireyi iyi tanıyan kişilerden öznel değerlendirmelerini istemek gibi başka değerlendirme yolları eklemek özellikle iyi bir fikirdir.

Dış Geçerlik Sorunları Tek denekli araştırma desenlerine sıklıkla yöneltilen bir eleştiri, bulguların dış geçerliğinin sınırlılığıdır. Diğer bir deyişle, tek denekli deneylerin vaka çalışması yöntemi ile aynı kısıtlılığı bulunmaktadır. Her birey benzersiz olduğu için, belirli bir müdahalenin etkisinin diğer bireylere genellenip genellenemeyeceğini bilmenin hiçbir yolu yoktur. Ancak tek denekli deney bulgularının dış geçerliğinin göründüğü kadar kısıtlı olmadığına ilişkin bazı nedenler vardır.

Birincisi, tek denekli deneylerde genellikle güçlü müdahaleler kullanılmaktadır ve bunlar dramatik ve büyük değişimler ortaya çıkarırlar (Kazdin, 1978). Buna bağlı olarak da bu tür müdahalelerin diğer bireylere de genellendikleri görülmüştür. Tek denekli deneylerin genellenebilirliğine ilişkin bir diğer kanıt da çoklu temel düzey desenlerinin kullanımından gelmektedir. Örneğin, bireyler arası çoklu temel düzey deseni, bir bireyde etkili olan bir müdahalenin diğer bireylerde de değişimi ortaya çıkardığını sıklıkla göstermektedir. Aynı şekilde, ortamlar ve davranışlar arasında yapılan çoklu temel düzey desenleri de müdahale etkisinin dış geçerliğine ilişkin kanıt oluşturmaktadır. İyi tanınan bir araştırmacının işaret ettiği gibi:

> Bilimde replikasyon (tekrar), bulguların güvenirliğini ortaya koyabilmek ve aynı zamanda da koşullar arası (denekler, araştırmacılar) geçerliği değerlendirmek için bir anahtardır. Biz birçok koşulda benzer sonucu olan, sağlam müdahaleleri belirlemek istemekteyiz. Müdahaleler koşullar arası dayanıklı olmadıklarında ise kısıtlılıklarını ve aracı değişkenleri bilmek isteriz. Bunun için birçok çalışma ve birçok çalışma yöntemi gereklidir. Onlarca yıllık endişeler ve araştırmalar sonrasında, tek denekli deneylerin, gruplararası deneylerin veya yarı deneysel çalışmalardan herhangi birisinin, bireyler veya yeni koşullar üzerinden daha iyi genellendiğine ilişkin net bir kanıt bulunamamıştır (Kazdin, 2001, s. 377).

Bunun psikoloji biliminde çok yöntemli bir yaklaşıma çağrı olduğunu görmek gerekir. Bir sonraki bölümde "yarı deneysel" yöntemleri tartışacağız.

ÖZET

Tek denekli araştırma desenlerinden önemli iki tanesi vaka çalışması ve tek denekli deneydir (küçük-*n* deseni). Vaka çalışması yöntemi, davranış hakkında hipotezler önerebilir, klinik müdahale için bir fırsat yaratabilir (örneğin, terapi için yeni yaklaşımları denemek gibi), ender olguların derinlemesine çalışılmasına izin verebilir, kuramsal varsayımlara meydan okuyabilir ve bir psikoloji kuramına geçici destek sağlayabilir. Vaka çalışması yönteminin karakteristik özelliği olan bireylerin yoğun incelenmesine idiyografik araştırma adı verilir ve psikolojiyi niteleyen nomotetik yaklaşıma (genel kurallar ve ilkeler arama) tamamlayıcı olarak görülebilir. Vaka çalışması yöntemi neden-sonuç kararına varmak için kullanılırsa, verinin toplanmasında veya yorumlanmasında yapılan yanlılıklar tanımlanmadığında, sorunlar ortaya çıkabilir. Vaka çalışması, tek bir bireyin bulgularının genellemesinde de potansiyel sorun barındırır. Dahası, bazı vaka çalışmalarında ortaya çıkan "dramatik" bulgular, bu yöntemin sınırlılığının farkında olmayan kişiler tarafından sıklıkla geçerli kabul edilir.

B. F. Skinner davranışın deneysel analizini geliştirmiştir. Uygulamalı davranış analizi, davranışın deneysel analizinden ortaya çıkan bilgiyi sosyal olarak ilişkili problemlere uygulamaktadır. Bu yaklaşımların temel yöntemi, tek denekli deneyler veya küçük-*n* desenleridir. Birçok tek denekli desen bulunmasına rağmen, en yaygın kullanılanları ABAB deseni ve çoklu temel düzey desenleridir.

ABAB deseni veya tersine desen araştırmacının, müdahalenin olmadığı (temel düzey) ve olduğu koşullar arasında davranışın sistematik olarak değiştiğini göstererek işlemin (müdahalenin) etkisini görmesini sağlar. Birinci işlem (B) aşamasında değişen davranış, ikinci temel düzey aşamasında (A) geri çekildiğinde tersine dönmez ise yöntemsel sorunlar ortaya çıkar. Böyle bir durumda, başlangıçtaki değişimin sorumlusunun müdahale mi yoksa başka bir etken mi olduğunu ortaya koymak zordur. ABAB deseninde, yararlı olduğu gösterilen bir müdahalenin geri çekilmesi de etik sorunlar doğurabilir.

Çoklu temel düzey deseni, müdahalenin bir sonucu olarak davranışın birden fazla temel düzeyde değiştiğini göstererek etkililiği kanıtlar. Başlangıçta farklı bireylerde veya aynı bireylerde farklı davranış ve ortamlarda temel düzeyler kaydedilir. Davranışın, müdahalenin sunumundan hemen sonra değişmediği veya işlem etkisinin diğer bireylere, diğer davranışlara veya diğer ortamlara genellendiği durumlarda yöntemsel sorunlar ortaya çıkmaktadır.

Temel düzeyde aşırı değişkenlik ve azalan veya artan temel düzeyler gibi problemler, tek denekli desenlerde bulguların yorumlanmasını zorlaştırmaktadır. Temel düzeyin aşırı değişken olması problemi değişkenliğin kaynağını arama, ortadan kaldırma ve temel düzey gözlem süresini arttırma yöntemleriyle aşılabilir. Artan veya azalan temel düzeylerin olması durumunda araştırmacı, müdahalenin etkililiğini gösterebilecek başka kanıtları aramaya ihtiyaç duyabilir. Son olarak tek denekli deneysel desenler sıklıkla dış geçerlik eksikliğinden dolayı eleştirilir. Ancak müdahaleler tipik olarak davranışta yeterli

miktarda değişimler gösterdiğinden, farklı bireylerde kolayca tekrarlanabilmektedir. Tek "grup" deneklerin kullanımı (küçük-*n* deseni) da denekler arası genellenebilirliğe hızlı bir kanıt sağlar. Psikoloji bilimine en iyi hizmet edecek yaklaşım, çok yöntemli bir yaklaşımdır ve replikasyon araştırma bulgularının kabulünde anahtar olmaya devam etmektedir.

TEMEL KAVRAMLAR

vaka çalışması 278
idiografik yaklaşım 284
nomotetik yaklaşım 284
tek denekli deney 290
temel düzey aşaması 290
ABAB deseni (tersine desen) 291
ortamlar arası çoklu temel düzey desenleri 294
bireyler arası çoklu temel düzey desenleri 296
davranışlar arası çoklu temel düzey desenleri 296

GÖZDEN GEÇİRME SORULARI

1. Vaka çalışması yöntemini tanımlayınız ve her bir avantajı için bir örnek veriniz.
2. İdiyografik ve nomotetik araştırma yaklaşımları arasındaki farkı anlatınız.
3. Vaka çalışması yöntemini tanımlayınız ve her sakıncası için bir örnek veriniz.
4. Neden-sonuç yargısına varmada vaka çalışması yönteminin temel kısıtlılığı nedir?
5. Hangi koşullarda tek denekli deneysel bir desen, çok gruplu bir desenden daha uygun olabilir?
6. Tek denekli deneysel bir desende, temel düzey ve müdahale aşamalarını ayırt ediniz.
7. ABAB deseni neden tersine döndürme deseni olarak adlandırılır?
8. ABAB deseni ile özellikle ilişkilendirilen yöntemsel sorunlar nelerdir?
9. Çoklu temel düzey desenlerinin tümünde ortak olan mantığı ve genel sürecin ana hatlarını anlatınız.
10. Özellikle çoklu temel düzey desenleriyle ilişkili olan yöntemsel sorunlar hangileridir?
11. Hangi yöntemsel sorunlar tüm tek denekli desenlerde ele alınmalıdır?
12. Hangi kanıtlar tek denekli desenlerin dış geçerliğini destekler?

DÜŞÜNDÜRÜCÜ SORULAR

1 Popüler bir dergide "çamur terapisinin" aşırı kaygılı bir bireyi tedavi etmede nasıl başarılı olduğunu anlatan bir vaka analizi bildirilmişti. Hastanın belirtileri uyku sorunu, iştah kaybı, grup içindeyken aşırı sinirlilik, kişinin kendini her zaman "diken üstünde" ve korkulu hissetmesine neden olan genel uyarılmışlık hislerini içermekteydi. Çamur terapisini uygulayan ve birkaç televizyon kanalında söyleşi programlarına çıkan Kaliforniyalı terapist, bu alanda tanınmış birisiydi. Terapist hastaya önce derin rahatlama tekniklerini ve onu rahatsız eden düşünceleri engelleyebilmek için tekrarlaması gereken bir "gizli sözcük" öğretti. Daha sonra hastadan günde iki saat boyunca çamur dolu özel bir tahta küvette, çamura batmış bir biçimde yatmasını istedi. Bu sürede hasta rahatlama alıştırmalarını

(devam ediyor)

çalışmakta ve en küçük bir kaygı yaşadığında gizli kelimeyi tekrarlamaktaydı. Terapi maliyeti yüksekti, ancak 6 hafta sonunda hasta daha önce şikâyet ettiği kaygılardan kurtulduğunu bildirmişti. Terapist onun iyileştiğini söylemiş ve tedavinin başarısını sakinleştirici çamura batmaya bağlamıştı. Magazin makalesinin yazarı terapiyi "birçok kişinin yararlanabileceği bir tedavi" olarak tanımlamıştır. Vaka çalışması yönteminin sınırlılıklarına ilişkin bilgilerinizden faydalanarak aşağıdaki soruları cevaplayınız:

- **A** Çalışmada olası hangi yanlılık kaynakları bulunmaktadır?
- **B** Başarılı tedaviye ilişkin hangi alternatif açıklamalar önerebilirsiniz?
- **C** Sadece tek bir bireyin incelenmesi muhtemel hangi sorunları getirmiş olabilir?

2 Beş yaşındaki bir çocukta sıklıkla cilt kızarıklıkları görülmektedir. Aile doktoru anneye problemin "çocuğun yediği bir şeyden" kaynaklandığını, bu nedenle çocuğun ne yediğine iyice dikkat etmesini söylemiştir. Anne çocuğun gün boyunca yediği yiyecekleri ve ertesi gün kızarıklık olup olmadığını kaydederek soruna yaklaşmaya karar vermiştir. Kızarıklığın varlığı ve yokluğu ile yiyecek arasında bir ilişki yakalamayı umut etmektedir. Bu yaklaşım belirli yiyecekleri yeme ile kızarıklık arasındaki ilişkinin bulunması için faydalı olabilecek olsa da, tek-denekli desenlerin mantığını ve işlemlerini kullanmak daha iyi bir yaklaşım olacaktır. Annenin böyle alternatif bir yaklaşımı nasıl kullanabileceğini anlatınız. Açık olunuz ve bu davranışsal yöntemin kullanımında ortaya çıkabilecek olası problemlere işaret ederek açıkça anlatınız.

3 Yaz aylarında hafif düzeyde zihinsel engelli çocukların bulunduğu bir kampta iş buldunuz. Danışman olarak küçük bir grup çocuğa bakmanız ve içeride gerçekleşen aktivitelerde (örneğin, elişi ve dikiş) dikkatlerini geliştirmek için yollar aramanız istenmektedir. "Görevle ilgilenmeye" bağlı bir ödül sistemini (örneğin, şeker) kullanma olasılığını incelemeye karar verdiniz. Kamp yöneticisinin müdahale stratejinizin etkililiği ve kamptaki diğer çocuklarda da işe yarayacağına ilişkin kanıt isteyeceğini de fark ettiniz. Bu nedenle aşağıdakileri soruları cevaplayınız.

- **A** Çocukların bir kamp faaliyetinde daha fazla zaman geçirmesini hedefleyen, pekiştirme ilkelerine dayalı bir müdahale stratejisi planlayınız.
- **B** Hangi davranışsal kayıtları tutmanız gerektiğini ve çocuğun davranışındaki değişime sizin müdahalenizin neden olduğunu nasıl belirleyeceğiniz anlatınız. Örneğin, davranışı ne zaman ve nasıl ölçeceğinizi açıklamanız ve bu "deneyinizi" neden bu özel desen ile yürüttüğünüzü gerekçelendirmeniz gerekecektir.
- **C** Bu müdahale stratejisinin diğer çocuklarda da işleyeceği konusunda yöneticiyi ikna etmek için kullanacağınız argümanları anlatınız.

4 Bir öğretmen, sınıfındaki problemli bir çocuğu idare edebilmeye yardımcı olabilecek bir davranışsal müdahale planlamak için sizden yardım istemektedir. Çocuk sırasında oturması istendiğinde oturmamakta, "sessiz zamanlarda" sessiz durmamakta ve öğrenme ortamını bozan başka davranışlar sergilemektedir. Şeker veya oyuncak gibi olumlu bir pekiştirecin, çocuğun davranışlarını düzeltmek amacıyla nasıl kullanılacağını, "davranışlar arası çoklu-temel düzey deseninin" bir parçası olabileceğini anlatınız..

İleri Araştırmaların Cevapları

1 Arkadaşınızla aynı fikirde olma eğiminde olabilirsiniz. Kişisel örnekler çoğunlukla nicel kanıtlardan daha ikna edicidir. Ancak bu iki çalışmayı değerlendirirken, ikisinin farklı araştırma yaklaşımlarını temsil ettiğini görmek önemlidir. Birincisi büyük gruplarla çalışmaya dayanan ve nicel ölçümler kullanmaya eğilimli olan nomotetik yaklaşıma örnektir. İkinci çalışma ise bireysel vakaların ve niteliksel betimlemelerin yapıldığı idiografik yaklaşımı temsil eder. Bu iki yaklaşımın arasındaki farklılıklar bilinirse ve bulgular dikkatlice incelenirse, bunlardan birisini seçmenin şart olmadığı görülür. Birinci çalışma evliliklerin yarısından biraz fazlasının boşanma ile sonuçlandığını gösterirken, yarısından biraz azının da boşanma ile sonuçlanmadığını göstermektedir. İkinci çalışmada ise anlaşmazlık ve aile geçmişi nedeniyle risk altında olan evliliklerin, mutlaka boşanmayla sonuçlanmadığı görülüyor. İkinci çalışma, bu risk faktörlerinin üstesinden gelebilmek için fazladan bir çaba gerekebileceğini düşündürmektedir. Örneğin, boşanmayı düşünen çift terapiye başladığı zaman, bir yıllarını evlilikleri üzerinde

çalışmak için terapide geçirmeye istekliydiler. Bu iki çalışmanın bulguları, nomotetik ve idiyografik çalışmaların, rakip değil birbirlerini tamamlayıcı olduğunu göstermektedir.

2 Arkadaşınızın ikinci sorusu, psikoloji öğrencilerinin genellikle sorduğu (ve sorması gereken) bir sorudur. Tüm bu kanıtların benimle ne ilgisi olabilir? Bu iki çalışmanın bulguları geleceğini düşünürken arkadaşınıza yararlı bir bilgi oluşturabilir. Birinci çalışma bize boşanmanın sık olduğunu ve boşanma olasılığının artmasına ilişkin bazı belirtilerin bulunduğunu göstermektedir. İkinci çalışma ise, evliliklerin bu etkenler olsa bile başarılabildiğini gösterir. Bu bilgi, deneyimlerden elde ettiğimiz bilgiyi sistematik ve kontrollü bir çalışmadan gelen kanıtlarla tamamladığı için yararlı olabilir. Arkadaşınız bu bulgulara dayanarak boşanıp boşanmayacağını bilemez veya buna bağlı olarak evlenip evlenmeyeceğine karar veremez. Genel olarak psikolojik araştırmaların bulguları bize, Gordon Allport'un sorduğu "Bu durumda, herhangi bir kişinin yapması gereken şey nedir?" sorusunun cevabını henüz veremiyor.

Düşündürücü Soruların Cevapları

A Bu vaka çalışmasındaki bir yanlılık kaynağı, aynı bireyin hem terapist hem de araştırmacı olarak görev yapması ve bunun gözlemci yanlılığı sorununa benzer problemleri ortaya çıkarmasıdır. İkinci bir yanlılık kaynağı ise, terapistin sonuçlarını sadece hastanın öz bildirimine dayandırmasıdır.

B Başarılı tedavi, sadece gevşeme tekniklerinin bir sonucu olabilir; kaygı karşısında "gizli sözcüğün" kullanımı, terapistin hastaya gösterdiği ilgi ve hatta tedavinin yüksek bedelinin sonucu olabilir.

C Dış geçerlik eksikliği, sadece tek bir bireyin incelenmesine ilişkin bir sorun olarak ortaya çıkabilir.

BÖLÜM ON

Yarı Deneysel Desenler ve Program Değerlendirme

BÖLÜMÜN ANA HATLARI

GENEL BAKIŞ
GERÇEK DENEYLER
Gerçek Deneylerin Özellikleri
Doğal Ortamda Gerçek Deney Yapmanın Zorlukları
Gerçek Deneylerde İç Geçerliğe Tehditler
Gerçek Deneylerde Bile Kontrol Edilemeyebilen Sorunlar
YARI DENEYSEL DESENLER
Denk Olmayan Kontrol Grup Deseni
Denk Olmayan Kontrol Grup Deseni: Langer ve Rodin'in Çalışması
Denk Olmayan Kontrol Grup Deseninde İç Geçerliğe Tehditler
Dış Geçerlik Meselesi
Kesintili Zaman Dizileri Deseni
Denk Olmayan Kontrol Gruplu Zaman Dizileri Deseni
PROGRAM DEĞERLENDİRME
ÖZET

GENEL BAKIŞ

En genel anlamıyla deney bir testtir ve henüz bilmediğimiz bir şeyi anlamak için kullandığımız bir işlemdir. Bu bağlamda bir yemek tarifine yeni malzemeler ekleyip lezzetinin artıp artmadığını görmek istediğimizde ya da iş yerimize daha hızlı varabilmek için farklı bir yol denediğimizde deney yapmış oluruz. Bilimsel yöntemin kullanılmadığı bu tip günlük deneyler, psikolojik araştırmalarda yürütülen deneylerden oldukça farklıdır. Gözlem ve anket gibi araştırma tekniklerinin aksine, deneysel yöntemler neden-sonuç ilişkisine karar vermede kullanılan en etkili yöntemler olarak kabul edilir. Fakat bir nedensellik olup olmadığına karar vermek kolay bir iş değildir ve son bir kaç bölümde bir olayı, onun nedenselliğini keşfederek anlamaya çalışan araştırmacıların karşılaştıkları karmaşık görevler ele alınmıştır.

Bu bölümde deneysel yöntemleri tartışmaya devam edeceğiz ama bu defa hastaneler, okullar ve iş yerleri gibi doğal ortamlarda yapılan deneyler üzerinde duracağız. Araştırmacılar, laboratuvarlarda elde edilen sonuçların dış geçerliğini test etmek ya da insanların yaşama ve çalışma koşullarını iyileştirmeyi amaçlayan bir tedavinin etkisini değerlendirmek için doğal ortamlarda araştırma yapabilir. Bu tip ortamlarda neden-sonuç çıkarımı yapmak daha zordur ve bir araştırmacı laboratuvar yerine doğal ortamlarda deney yapmayı tercih ettiğinde yeni problemler ortaya çıkar.

Birçok psikoloğun araştırma alanları, onları laboratuvardan çıkıp alanda çalışmaya sevk eder. Laboratuvarlarda çalışan araştırmacılar sıklıkla bir olayı anlamak veya bir teori test etmek gibi amaçlarla *temel araştırmaya* odaklanır. Bunun aksine, doğal ortamlarda *uygulamalı araştırma* yapan psikologların genellikle gerçek hayat koşullarını iyileştirmeye yönelik pratik amaçları vardır. Fakat bununla birlikte, doğal ortamlarda araştırma yapanlar, laboratuvar koşullarındaki üst düzey bilimsel kontrolü çoğunlukla kaybederler. Doğal ortamlarda yürütülen "gerçek dünya deneyleri" karmaşık olabilir. Örneğin bir iş yerinde bağımsız değişkenin çeşitli koşullarına atanan katılımcılar, iş değişikliği ya da şirket uygulaması nedeniyle deneyden ayrılabilir ve bu durum deney süresince kaydedilen bağımlı değişkeni birçok şekilde bozar. Doğal ortamlarda yürütülen araştırmaların değerlendirilmesinde önemli bir bölüm ise çalışma boyunca deneysel kontrolün ne derecede sağlandığını incelemektir.

Laboratuvar ortamlarında mevcut yüksek derecedeki kontrol, araştırmanın iç geçerliğini (nedensel çıkarım yapabilme yeteneğini) güçlendirir, ama bu avantajın bir de bedeli vardır: Bu yapay laboratuvar ortamı sonuçların dış geçerliğini (sonuçları, araştırmanın ortamının dışında kalan insanlara, ortamlara ve koşullara genelleyebilme yeteneği) düşürür. Belli bir olay laboratuvar koşullarında araştırıldıktan sonra araştırmacılar, deneylerini doğal ortamlarda yürüterek sonuçların dış geçerliğini saptamak isteyebilir. Uygulamalı bir araştırmanın yüksek dış geçerliği, özellikle geniş ölçekli sosyal değişimlerin hesaba katıldığı zamanlarda çok önemli olabilir.

Doğal ortamlardaki deneylerin halk ve toplum üzerinde geniş kapsamlı bir etkisi olabilir. Örneğin, Susam Sokağı (Sesame Street) televizyon dizisi ve engelli çocuklar için "Avantaj" (the Head Start) programı, yüz binlerce çocuğun eğitimini geliştirmek için tasarlanmış sosyal deneylerdi (bk. Şekil 10.1). Laboratuvar deneyleri sadece araştırmaya katılan bir kaç araştırmacı ve katılımcıyı

ŞEKİL 10.1 Susam Sokağı sosyal bir deney olarak yüz binlerce çocuğun eğitimini geliştirmek için tasarlandı.



etkilerken, okul, iş yerleri ve hastaneler gibi doğal ortamları kapsayan küçük çaplı deneylerin bile deneye dâhil olan insanlar üstünde geniş ve anlık etkileri olabilir. Belli bir tedavinin doğal bir ortamda etkili olup olmadığını anlamak için tasarlanan bir araştırma, tedavinin uygulanması için ek para ve zaman ayrılması gerekip gerekmediği ile ilgili olan kararı etkiler. *Programı değerlendirmenin* amacı, enstitüler, devlet kurumları ve diğer kuruluşlar tarafından yapılmış değişikliklerin ne derecede etkili olduğuna karar vermektir.

Bu bölümde doğal ortamlarda deney yaparken karşılaşılan engelleri anlatıp, üstesinden nasıl gelineceğini açıklayacağız ki mümkün olan her durumda gerçek deneyler yapılabilsin. Bununla birlikte, gerçek deneyler bazen laboratuvar dışında verimli olmaz. Bu gibi durumlarda, ancak laboratuvar koşullarına yakın deneysel işlemler düşünülmelidir. Bu *yarı deneysel* tekniklerin çeşitli şekillerini ele alacağız. Ve program geliştirmenin mantığına, işlemlerine ve sınırlılıklarına kısa bir giriş yaparak bölümü sonlandıracağız.

GERÇEK DENEYLER

Gerçek Deneylerin Özellikleri

Gerçek deneylerde araştırmacılar, bağımsız değişkeni değişimler, koşulları karşılaştırır ve (özellikle koşullara seçkisiz atama yaparak) üst düzeyde kontrol uygular.

Birçok günlük aktivitemiz aslında birer deney olabilmekle birlikte (bir yemek tarifinin malzemelerini değiştirmek gibi), bu tip deneyleri kitap boyunca ele alacağımız "gerçek" deneyler gibi düşünemeyiz. Bir analoji olarak hükümet, şirket yetkililerinin veya eğitim yöneticilerinin yürüttükleri "sosyal deneyler" de gerçek deneyler değildir. *Gerçek bir deney, bir olayı meydana getiren belirli bir neden ortaya çıkaran deneydir.*

Gerçek deneylerin üç önemli özelliği vardır:

1. Gerçek deneylerde bir çeşit müdahale ya da uygulama yapılır.
2. Gerçek deneylerin belirleyici özelliği, deneycinin şu hususlar üzerinde gösterdiği yüksek dereceli kontroldür: Deneysel koşulların düzenlenmesi, katılımcıların atanması, bağımsız değişkenlerin sistemli şekilde değişimlenmesi ve bağımlı değişkenlerin seçimi. Katılımcıların deneysel koşullara seçkisiz olarak atanması ise sıklıkla gerçek bir deneyin en önemli özelliği olarak görülür (West & Thoemmes, 2010).
3. Son olarak, gerçek deneylerde deneyciler, uygulamanın etkiliğini değerlendirmek için uygun bir karşılaştırma şekli belirler. En basit deneysel durumlarda bu karşılaştırma, iki denk grubun incelenen değişken dışında tamamen aynı etkenlere maruz kalmasıdır.

Gerçek deneyin koşulları karşılandığı takdirde, bağımlı değişkende gözlenen herhangi bir fark, mantıksal olarak bağımsız değişkenin düzeyleri arasındaki farka bağlanabilir. Bununla birlikte, doğal ortamlarda yapılan gerçek deney ile laboratuvarlarda yapılanlar arasında farklılık vardır.

Doğal Ortamda Gerçek Deney Yapmanın Zorlukları

- Araştırmacılar doğal ortamlarda gerçek deney yürütebilmek için izin alma ve katılımcılara ulaşma konusunda zorluk çekebilirler.
- Seçkisiz atama, kişileri yeni bir uygulamadan mahrum bıraktığı gerekçesiyle kimilerince haksız bir yöntem gibi algılanmasına rağmen, hâlâ bir uygulamanın etkili olup olmadığına karar vermek için kullanılan en adil ve iyi yöntemdir.

Deneysel yöntem, problem çözme ve uygulamalı sorulara cevap vermede kullanılan etkili bir araçtır. Yine de doğal ortamlarda deney yürütmeye çalışırken iki büyük engel ortaya çıkar. İlki, deneyi yürütmek için yetkililerden izin almaktır. Eğitim müdürleri, hükûmet ve iş adamları araştırmanın gerçekten faydalı olacağına ikna olmadan finansal ya da başka herhangi bir şekilde araştırmayı desteklemeye gönüllü olmaz. Doğal ortamlarda deney yapmanın bir ikinci ve daha büyük olan engeli ise katılımcılara ulaşma problemidir. Bu problem özellikle katılımcıların müdahale ya da denk gruba seçkisiz atanması söz konusuysa daha da sıkıntılı bir hâl alır.

Kaynakların sınırlı olduğu ve uygulamayı isteyenlerin hepsine uygulanamadığı bazı durumlarda seçkisiz atama kullanılabilir. Böyle bir durum 2008 senesinde Oregon'da yaşandı; devlet, düşük gelirli 10.000 kişiye daha sağlık sigortası sağladı (Klein, 2013). Bu iyi haberin bir de olumsuz yanı vardı: Sigortaya

ihtiyacı olanların sayısı 90.000 kişiydi. Devlet, insanların bu haktan eşit (seçkisiz) yararlanabilmesi için kura çekimi yaptı ve insanlar koşullara yansız olarak atandılar: Sağlık sigortasından faydalanacak olanlar ve faydalanamayacak olanlar (kontrol grubu). Bu deneyin sonuçlarına göre, sağlık sigortası sahibi olmak sağlık hizmetlerini daha etkili hâle getirdi, faydalanma oranını arttırdı, şeker hastalığı tanısını ve kontrolünü geliştirdi, depresyon oranını düşürdü ve bu haktan faydalananların maddî sıkıntılarını programın ilk iki yılında azalttı (Baicker ve ark., 2013).

Koşullara seçkisiz atama ilk bakışta haksız bir yöntemmiş gibi görünüyor – en nihayet, sağlık sigortası gibi faydalı bir uygulamadan bazı katılımcılar mahrum kaldı. Seçkisiz atamanın ne ölçüde hakkaniyetli olduğu, daha küçük çaplı araştırmalarda da bir sorundur. Varsayın ki araştırmacılar bir üniversitede alkol sorunuyla mücadele eden öğrencilere yardım etmek için hazırlanmış yeni bir tedaviyi test etmek istiyor. Araştırma yöntemleri bilginiz size yeni uygulamanın etkililiğine karar vermek için eski bir uygulamayla (ya da denk bir koşulla) karşılaştırılması gerektiğini söyler. Birbirine denk öğrenci grupları oluşturmak için seçkisiz atama kullanılmalıdır. Fakat seçkisiz atama adil görünmüyor. Kontrol grubuna atanan öğrenciler yeni uygulamanın faydalarından mahrum kalır. Bütün öğrencilerin faydalanabileceği yeni bir uygulama daha adil olmaz mı?

Seçkisiz atamanın ne kadar adil olduğuna bakalım. Eğer araştırmacılar bu yeni uygulamanın eskisinden daha etkili olduğunu bilselerdi, bu araştırmayı yapmaya ve yeni uygulamanın herkese verilmesine gerek kalmazdı. Bununla birlikte, eğer yeni uygulamanın etkili olup olmadığını bilmiyorsak, doğru bir deneye ihtiyaç vardır. Koşullara seçkisiz atama, öğrencileri farklı gruplara ayırmak için kullanılabilecek en adil yöntemdir. En nihayetinde, yeni uygulama eski bir uygulamadan *daha az* etkili bile olabilir.

Bu yaklaşımı kabul eden bir çalışma kapsamında büyük bir üniversitenin sağlık merkezinden tıbbî destek alan öğrenciler, sağlık geçmişleri ve davranışları ile ilgili kısa bir anket doldurdu (Schaus ve ark., 2010). "Alkol tüketiminde 5/4" düzeyindeki öğrenciler yüksek risk grubu olarak araştırmaya alındılar. Erkekler için yüksek riskte olma ölçütü, önceki iki hafta içinde bir seferde beş ya da daha fazla, kadınlar için ise dört ya da daha fazla alkollü içecek tüketimiydi. Çalışmaya katılmayı kabul eden yüksek riskteki öğrenciler (N = 363) her iki koşuldan birine seçkisiz olarak atandı. Uygulama grubunda olan öğrenciler, bilişsel-davranışsal beceriyi, motive edici strateji ve geri bildirimle birlikte kullanan iki motive edici kısa müdahale (MKM) aldı. Aynı zamanda bu öğrencilere, alkolden uzak durma konusunda bilgi verilen birer broşür dağıtıldı. Kontrol grubundaki öğrencilere ise müdahale edilmedi, sadece broşür verildi. Bütün katılımcılardan bu bir senelik çalışma boyunca anket doldurmaları, alkol kullanma konusunda günlük tutmaları ve bu bir yıl boyunca kandaki alkolü belirleyen testler istendi. Sonuçlara göre, müdahale gören öğrenciler kontrol grubundakilere göre, ortalama olarak daha az alkol tükettiler, risk düzeylerini düşürdüler ve alkolle bağlantılı sorunları daha az yaşadılar.

Çalışmanın bahsedilmeye değer bir başka yönü de şudur: Alkol tüketimi ile ilgili soruşturma sonunda çalışma için uygun olan 2000 öğrencinin

%70'i kendileriyle iletişime geçilmesi ve iki koşuldan birisine seçkisiz olarak atanmasını uygun bulmadı. Yeni tıbbî uygulamanın test edileceği klinik denemelerde, hastaları seçkisiz olarak gruplardan birisine ikna etmek çok zor olabilir. Tıbbî müdahale isteyen insanlar genellikle kontrol gruplarına atanmak istemezler.

Bir tedaviye (uygulamaya) olan talep eğer karşılanabilecek olandan fazlaysa, uygulama isteyen kişiler *bekleme listesi kontrol grubu* hâline getirilebilir. Bu kişiler uygulamayı beklerken araştırma anketlerini doldurur. Bununla birlikte, ilk gelenler uygulama grubuna yazılmaz; seçkisiz olarak uygulama grubuna veya bekleme listesi kontrol grubuna atanır. İlk gelen katılımcılar, son gelenlerden önemli ölçülerde farklılık gösterir (örneğin uygulama almak için daha istekli olmak gibi). Bu kişisel farklılıklar, deney ve kontrol grubu arasında, araştırmanın sonunda gözlenebilecek farklar için alternatif açıklamalar hâline gelir. Seçkisiz atama, bu tip özellikleri gruplara yansız ve eşit şekilde dağıtmak için gereklidir.

Hem denk gruplar oluşturmak hem de bütün katılımcılara uygulama sağlayabilmek için kullanılan bir diğer alternatif yöntem ise uygulamaları seçenekli yapmaktır. Örneğin, dördüncü sınıf öğrencileri derste elektronik tablet kullanmanın, materyalin anlama düzeyini etkileyip etkilemeyeceğini anlamak için test edilebilir. Uygulama (elektronik tablet), dil sanatları ve bilime giriş derslerinde yapılabilir, fakat elektronik tablet kullanacak öğrenciler bu iki konudan birine seçkisiz olarak atanır. Dil sanatları dersinde elektronik tablet kullanan öğrenciler, bilime giriş dersinde tablet kullanan öğrencilerin, bilime giriş dersi öğrencileri ise dil sanatları öğrencilerinin kontrol grubu olur. Deney bittikten sonra öğrencilere elektronik tabletle çalışamadıkları diğer konuyla ilgili çalışma fırsatı verilir. Bu şekilde bütün öğrenciler hem dil sanatları hem de bilime giriş dersinde uygulamadan faydalanmış olur.

Seçkisiz atamanın kullanılamayacağı durumlar her zaman olacaktır. Bu durumlarda araştırmacılar kendi araştırma sorularını cevaplamak için *yarı deneysel desenlere* yönelirler. Yarı deneysel desenlerin mantığı ve yönteminden bölümün ilerleyen kısımlarında bahsedilecektir.

Gerçek Deneylerde İç Geçerliğe Tehditler

- İç geçerliğe yönelik tehditler, araştırma sonuçlarının olası açıklamaları olabilecek karıştırıcılardır.
- İç geçerliğe yönelik başlıca tehdit şekilleri tarih, olgunlaşma, test etkisi, araç düzeni, ortalamaya gerileme, katılımcı kaybı, seçim ve arttırıcı etkidir.

Özellikle belirttik ki kontrol teknikleri bilim insanlarına, alternatif açıklamaları eleme olanağı sağlar; böylece, bağımlı değişkende gözlenen farkların nedeninin bağımsız değişken olduğu öne sürülebilir. Önceki bölümlerde karıştırıcı olarak adlandırılan bu alternatif açıklamalar, çalışmanın iç geçerliğini tehdit eder. Araştırmacılar, doğal ortamlarda kontrol tekniklerini genelde daha az kullanabildikleri için, yarı-deneysel desenlerde uzman olan araştırmacılar, karşılaşılabilecek başlıca karıştırıcı tiplerini belirlediler (Shadish, Cook & Campbell,

Temel Kavram

2002; West, 2010). **İç geçerliğe tehdit** olarak kabul edilen sekiz tip karıştırıcı vardır. Bunların bir kısmına zaten aşinasınız, diğerleri yenidir. Bu temel iç geçerliğe tehdit tiplerini gözden geçirdikten sonra, hangi yarı-deneysel işlemlerin bunları ne derecede kontrol etmeye çalıştıklarını inceleyeceğiz.

Tarih Etkisi Bir olayın meydana gelmesi, uygulamanın dışında, katılımcıların davranışında değişikliklere yol açıyorsa iç geçerliği tehdit edebilir. Gerçek bir deney, deney grubu ve kontrol grubu katılımcılarına uygulama dışında eşit şekilde davranılmasını (deney esnasında aynı uygulamalara tâbi tutulmalarını) gerektirir. Laboratuvar ortamında bu, koşullar aynı tutularak veya dengelenerek yapılır. Bununla birlikte doğal ortamlarda deney yapılırken, araştırmacı yüksek düzeyde bir kontrol sağlayamayabilir, bu yüzden tarih etkisine (deneyin geçmişine) bağlı olarak ortaya çıkan bozucu etki, iç geçerliği tehdit edebilir. Örneğin, mahalli polisin hırsızlık ve diğer genel rahatsızlık verici olaylara karşı yeni bir "mahalle izleme" programını test ettiğini varsayın. Programın etkililiğini ölçmek için polis kayıtları kullanılabilir. Bununla birlikte, uygun bir karşılaştırma grubu kullanılmadan, uygulamanın (mahalle izleme programının) dışındaki bazı başka olaylar suç oranını düşürürse, **tarih** (deneyin geçmişi) çalışmanın iç geçerliğini tehdit eder. Örneğin, önde gelen bir işletmenin bir hırsızlığa uğradığını ve medyanın büyük ilgisini çektiğini düşünün. Bu mahallenin bu haberi de kapsayan tarihi, uygulamanın etkisini belirlemeyi engelleyen bir karıştırıcı etki hâline gelir ve araştırmanın iç geçerliğini tehdit eder.

Temel Kavram

Olgunlaşma Deneydeki katılımcılar zamanla ister istemez değişirler, yaşlanırlar, daha deneyimli hâle gelirler ve buna benzer bir çok değişim olur. Tek başına zamanın akışıyla oluşan değişim olgunlaşma olarak adlandırılır. Örneğin, okula başlayan evsiz çocuklar için düzenlenen kahvaltı programının, öğrenim yılı boyunca akademik performanslarını etkileyip etkilemeyeceğini test eden bir okul psikoloğu düşünün. Uygun bir karşılaştırma olmadan bu araştırmacı, çocukların okulun başından sonuna kadar gösterdikleri performans değişikliğini kahvaltı programının etkisine bağlayabilir. Oysa gerçekte bu değişim sadece araştırmanın geçerliğini tehdit eden **olgunlaşma** yüzündendir. Çocukların öğrenme performansındaki ilerlemenin nedeni, sadece onların yaşla birlikte gelişen bilişsel becerileri olabilir.

Temel Kavram

Test Etkisi Bir test almış olmak, genelde sonradan alınan testlerdeki performansı etkiler. Örneğin, birçok öğrenciye bir derste verilen iki testten ikincisinin performansının, ilk testtekinden yüksek olduğu gerçeğini düşünün. Öğrenciler ilk testte, test işlemi ve öğretmenin beklentisi hakkında bir aşinalık kazanırlar. Bu aşinalık, onların ikinci testteki performansını etkiler. Aynı şekilde, birden fazla testin verildiği bir psikoloji deneyinde (örneğin, ön test-son test deseninde), **sınama etkisi**, uygulamanın etkisinden ayrılamazsa, bu durum iç geçerliğe bir tehdit oluşturur.

Temel Kavram

Araç Düzeni Sadece katılımcılarda (örneğin; olgunlaşma ya da test etkisi) değil, katılımcıların performansını ölçmek için kullanılan araçlarda da zamanla değişiklikler olabilir. Bu en çok, araştırmacının davranış

değerlendirmeye alışkın olduğu durumlarda ortaya çıkar. Örneğin, gözlemci yanlılığının nedenleri yorgunluk, beklenti ve gözlemcinin diğer özellikleri olabilir. Kontrol edilmediği sürece gözlemcilerdeki bu değişiklikler, gözlemler arasındaki davranış farklılıkları için alternatif açıklamalar getirerek iç geçerliğe yönelik **araç düzeni** tehdidi oluşturur. Davranışı ölçmek için kullanılan mekanik araçların kullanım sayısı arttıkça değişebilir ve devletin düzenlemeleri gözlemlerin yapılma şeklini değiştirebilir. Örneğin, eğitim programlarının ve hizmetlerinin öğrencilerin akademik performansı üzerindeki etkisini araştıran araştırmacılar, bağımlı değişken olarak çoğunlukla standart test sonuçlarını kullanır. 2014 yılında yeni standart testler hazırlandı; (www.corestandards.org), öğrencilerin farklı zamanlardaki performansını çalışan araştırmacılar, sadece standart testler değiştiği için bile performansın değiştiğini görecektir.

Temel Kavram

Temel Kavram

Ortalamaya Gerileme (Regresyon) İstatistiksel **regresyon**, katılımcıların puanlarının "aşırı" olduğu her durumda bir sorun teşkil eder. Bir testte aşırı olan puanlar, bir başka testte o kadar aşırı olmayabilir. Bir başka ifadeyle, çok çok iyi ya da çok çok kötü bir performanstan sonra (ikisine de sıkça rastlanır) o kadar iyi ya da o kadar kötü olmayan bir performans gözlenebilir. Örneğin sınıftaki en iyi test puanınızı düşünün. Bu testteki başarınızın nedeni neydi? Hiç şüphe yok ki iyi çalışmanızdı. Fakat şansınızın yaver gitmiş olması da mümkündür. Çok yüksek bir performans göstermek için her şeyin doğru şekilde işlemesi gerekir. Eğer bir sınavdan bahsediyorsak, sınavda çıkan konular en çok çalıştığınız yerden ya da hoşlandığınız bir test formatında çıkmış olabilir ya da kendinize özellikle güven duyduğunuz bir zamanda yapılmış olabilir veya bunların birçoğunun ya da hepsinin birlikte ortaya çıkmasından olmuş olabilir. İyi performanslar özellikle "aşırı"dır çünkü şansla (olağan ya da normal performansımızın üstünde) yükselmişlerdir. Aynı şekilde, kötü bir test performansının nedeni de şanssızlık olabilir. Tekrar test edildiğimiz zaman (çok iyi ya da çok kötü bir performanstan sonra), şansla ilgili etkenler bizi o en iyi ya da en kötü puana ulaştıracak şekilde tekrar "bir araya toplanmayabilir". Muhtemelen farklı zamanlarda aldığımız tüm puanlarımızın ortalamasına yakın bir puan alırız. Bu olay çoğunlukla, *ortalamaya gerileme* olarak adlandırılır (Kutu 10.1). İstatistiksel regresyona, test ya da ölçümlerin güvenilir olmadığı zamanlarda daha çok rastlanır. Kullanılan test güvenilir olmadığı zaman, skorların zaman içinde tutarsız olmasını bekleriz.

Şimdi okulun ilk döneminin sonunda (ön test) çok zayıf bir akademik performans göstermiş bir grup öğrencinin, bu performanslarını yükseltmeye çalıştıklarını düşünün. Bu öğrenciler, uç derecede düşük puan aldıkları için seçildi. Daha sonra bir uygulama (örneğin, 10 saatlik bir çalışma becerisi semineri) yapıldığını varsayalım. İstatistiksel regresyon iç geçerliğe yönelik bir tehdittir, çünkü bu *herhangi bir uygulama olmaksızın* sadece istatistiksel regresyona bağlı olarak bu öğrencilerin ikinci dönemin sonunda (son test) nispeten daha iyi bir performans göstermelerini bekleyebiliriz. Bu konuda bilgisi olmayan bir araştırmacı yanlışlıkla bu "regresyon etkisi"ni "uygulama etkisi" (seminerin etkisi) ile karıştırabilir.

KUTU 10.1

"HABER OLMA" UĞURSUZLUĞUNU AÇIKLAMA

Ortalamaya doğru gerilemenin iyi bilinen bir örneği "*Haber olma Uğursuzluğu*"dur. Bu iddiaya göre eğer bir atlet *Sports Illustrated (SI)* dergisinin kapağında haber olursa, bir sonraki yarışında bu bir uğursuzluk getirir ve atlet başarısız olur. Dergi çalışanları atletlerin performanslarını düşürecek bir büyü mü yaparlar?

İnsanların olaylar arasında nedensel bir ilişki kurmada özel bir yetenekleri vardır. Bununla birlikte bu eğilim çoğunlukla onları yanlı bir düşünceye sevk eder (Kahneman, 2011). İnsanların *haber olma uğursuzluğunu* açıklamak için öne sürdükleri bazı nedenler şunlardır.

- SI dergisinin kapağına çıktıktan sonra atletlerden beklentilerin yükselmesi ve üzerlerindeki baskının artması;
- SI dergisinin kapağına çıktıktan sonra atletlerin kendilerine çok güvenli ve kibirli hâle gelmesi;
- Diğer atletlerin, dergi kapağında çıkan atlete karşı çok daha sıkı çalışması.

Bu nedensel açıklamalar makul görünmelerine rağmen, hiçbirisi dergi kapağında boy gösterdikten sonra "uğursuzluk" sebebiyle atletlerin zayıf performans sergilemelerini *ortalamaya doğru regresyon* ile açıklamaya gerek duymaz.

İşte açıklaması: Bir atletin *SI* dergisinin kapağında görünmesi için alışılagelmişin dışında iyi bir performans göstermesi gerekir. Bu derece iyi bir performans (herhangi bir alanda), yetenek, ve *iyi şansın* kombinasyonudur. İstatistiksel olarak, aynı şansın tekrar ortaya çıkması pek olası değildir (şans, niteliği gereği, rastlantısaldır), bu yüzden bir sonraki performans muhtemelen o kadar büyük olmayacaktır. Sonraki performans ortalamaya yakın olur; bir başka deyişle, performans ortalamaya doğru geriler. Aynı istatistiksel muhakeme, bir sezonda şampiyon olan takımların bir sonrakinde de şampiyon olmakta neden güçlük çektiğini anlamak için de yapılabilir. Şampiyonluğun kazanıldığı sene yakalanan şansın gelecek senede de tekrarı olasılığı düşüktür.

Ortalamaya doğru regresyonu test etmek için, üstün performansları *SI* dergisinin kapağına taşınmamış olan atletleri spor veri tabanlarından araştırmayı deneyebilirsiniz (karşılaştırma grubu). Onların sonraki performanslarının da ortalamaya yakın olduğunu göreceksiniz. İnsanların istatistiksel açıklamaları kabul etmeleri zordur; aynı şekilde, performansı iyi ya da kötü şansa bağlamayı da sevmeyiz. Fakat "*SI* Uğursuzluğu" gibi durumlar söz konusu olduğunda, sadece ortalamaya doğru regresyon açıklaması ihtiyacımızı karşılıyor.

Temel Kavram

Katılımcı Kaybı Altıncı Bölüm'de bahsedildiği gibi katılımcılar deneyden ayrıldıkları zaman iç geçerliğe yönelik bir tehdit oluşur (örneğin, araştırma projesinden çekildikleri zaman). İç geçerliğe tehdit olan **katılımcı kaybı**, kaybedilen katılımcıların uygulama için deneyden önce oluşturulmuş grubun niteliğini değiştirdikleri varsayımına dayanır –örneğin, seçkisiz atama yoluyla oluşturulmuş grupların eşitlikleri bozulur. Bu durum, uygulama nahoş ise ve deney grubundaki bazı katılımcılar hayal kırıklığına uğrayıp, deneyi bıraktıkları zaman olabilir. Deney grubunda kalan katılımcılar, deneyi bırakan katılımcılardan (ve kontrol grubundakilerden) farklılaşır, çünkü bu onların muhtemelen uygulamayı daha çok istediklerini gösterir.

Temel Kavram

Seçim Çalışmanın başlangıcından itibaren gruplardaki bireylerin özellikleri arasında farklılıklar olduğu zaman, iç geçerliğe yönelik **seçim** tehdidi oluşur. Bu, uygulama grubundaki kişilerin, karşılaştırma grubundaki insanlardan, grup atanması dışında (uygulama alıyor olmaları dışında) birçok şekilde farklılaştığı anlamına gelir. Laboratuvarda, iç geçerliğe yönelik bu tip bir tehdit genellikle, katılımcıları seçkisiz atama yoluyla dengeleyerek ortadan kaldırılır.

Çalışmalarını doğal ortamlarda yürüten araştırmacılar, katılımcılarını çoğu kez deney ya da kontrol grubuna seçkisiz olarak atayamazlar; bunun sonucu olarak seçim, iç geçerliğe olası bir tehdit hâline gelir.

Seçimle Gelen Birleşik Etkiler Tarih ve olgunlaşma gibi iç geçerliğe yönelik kişisel tehditler ek bir endişe kaynağı olabilir, çünkü seçim ile birleşip iç geçerliğe yönelik tehdit oluşturabilirler. Özellikle denk gruplar seçkisiz atama yoluyla oluşturulmadığı zaman, (1) seçim ve olgunlaşma, (2), seçim ve tarih ve (3) seçim ve araç düzeninin birleşik etkilerine bağlı olan problemler ortaya çıkar. Örneğin, eğer bir okulda deney grubu olarak kullanılan birinci sınıf öğrenciler, ikinci sınıf öğrencileri ile karşılaştırılırsa, *seçim ve olgunlaşmanın birleşik etkileri* oluşabilir. Öğrencilerin ilk yılda yaşadıkları değişimlerin (okul çevresine aşinalık kazanırken), ikinci sene yaşadıkları değişimlerden çok daha fazla olduğu varsayılabilir. Deney ve kontrol grubu arasında gözlenen farklılıkların nedeni, deneysel değişimleme değil, olgunlaşma oranlarındaki bu farklılıklar olabilir.

Seçim ve tarihin birleşik etkileri, belirli bir zamanda meydana gelen olayların bir grup katılımcıda, diğer bir grup katılımcıdan farklı bir etki yarattığı zaman ortaya çıkar. Bu durum özellikle gruplar bir bütün olarak (intact grups) karşılaştırıldığında problemdir. Mesela sadece belli bir gruba has olaylar olduğu zaman, o olayın o grup üstündeki etkisi diğer gruplarınkinden çok daha fazla olur. Örneğin, iki okul kampüsünü kapsayan (bir uygulama, bir de kontrol grubu) bir AIDS farkındalığı kampanyasının etkisini araştıran bir çalışma düşünün. Makul şekilde, ülke çapında AIDS'e karşı gösterilen basın ilgisinin her iki kampüsteki öğrencileri eşit derecede etkilediğini farz edebiliriz. Bununla birlikte eğer bir okuldaki öğrenci AIDS sebebiyle ölmüş ve bu haber okul gazetesinde yer almışsa, bu okuldaki katılımcıların diğer okuldaki katılımcılardan çok daha farklı bir şekilde etkileneceğini varsayabiliriz. AIDS farkındalığı kampanyasının etkilerini ölçmek açısından, bu durum seçim ve tarihin birleşik etkisini temsil eder.

Son olarak, *seçim ve araç düzeninin birleşik etkisi*, test aracının bir gruptaki değişiklikleri diğerinden nispeten daha hassas ölçmesiyle ortaya çıkar. Örneğin, bu durum taban veya tavan etkisi olduğunda gözlenir. Bir grup, bir ölçümden çok düşük bir puan aldığı (taban etkisi) ve grup içindeki çok düşük puanlar güvenilir bir şekilde ölçülemediği zaman ya da aynı şekilde grup puanları çok yüksek olduğu zaman (tavan etkisi) durum böyledir. Kolayca anlaşılacağı gibi, eğer deneysel grup taban ya da tavan etkisi nedeniyle nispeten hiçbir değişim göstermezken, kontrol grubu güvenilir bir düzeyde değişim gösterirse (çünkü performans ortalamaları zaten başından ölçeğin ortasına yakındır) iç geçerliğe yönelik tehdit oluşur.

Gerçek deneylerin en büyük avantajlarından birisi iç geçerliğe yönelik olan bütün bu tehditleri kontrol etmeleridir. Campbell'in (1969) vurguladığı gibi, mümkün olduğu sürece gerçek deneyler yapılmalıdır; fakat bu mümkün olmadığı zamanlarda, yarı deneysel desenler uygulanmalıdır. "Elimizdekilerle yapabileceğimizin en iyisini yapmalıyız" (s. 411). Yarı deneysel desenler, bir uygulamanın etkililiğine dair geçerli bir bilgi kazanma amacı ve gerçek deneyler yürütmenin her zaman mümkün olmadığı farkındalığı arasındaki en elverişli uyumu temsil eder.

Gerçek Deneylerde Bile Kontrol Edilemeyen Sorunlar

- Herhangi bir araştırmada ortaya çıkabilecek iç geçerliğe yönelik tehditler kirlenme, deneyci beklentisi etkisi ve yenilik etkisini kapsar.
- Deney hakkındaki bilgi, gruplar arasında katılımcılar tarafından birbirlerine aktarıldığı zaman kirlenme oluşur. Bu durum dargınlık, rekabet ya da uygulamanın yayılmasına yol açar.
- Deneycinin yanlılığı ve beklentileri, kasıt olmaksızın, araştırma sonuçlarını etkilediğinde deneyci beklenti etkisi ortaya çıkar.
- Hawthorne etkisini de kapsayan yenilik etkileri, insanların davranışlarının sadece herhangi bir yeniliğin (örneğin uygulama) yol açtığı heyecan, enerji ve heves nedeniyle değişmesiyle oluşur.
- Dış geçerliğe yönelik tehditler, bağımsız değişkenin etkisi belirli insan, ortam, uygulama ve deneyin dışına genellenemediği zaman ortaya çıkar.

Belirli yarıdeneysel işlemler üzerinde düşünmeden önce, gerçek deneylerin bile deneysel sonuçların yorumu sırasında oluşabilecek olası tehditlerin hepsini kontrol edemeyeceğini göstermeliyiz. Gerçek deney yaparak iç geçerliğe yönelik büyük tehditleri ortadan kaldırabilmemize rağmen, doğal ortamlarda çalışan araştırmacılar ek bazı tehditlere karşı korunmak zorundadır. Bu tehditler kirlenme, deneyci beklentisi etkisi ve yenilik etkisidir.

Temel Kavram

Kirlenme, gruplardaki katılımcılar arasında deney hakkında iletişim olduğu zaman ortaya çıkar. Gruplar deney hakkında bilgi alış verişi yaparken istenmeyen birçok etki oluşur (Cook & Campbell, 1979). *Dargınlık* ise kontrol grubundaki kişilerin, arzu edilir bir uygulamanın kendilerine yapılmadığını öğrendikleri zaman ortaya çıkabilir. Örneğin, bir işyerinde performans arttırıcı hususları test eden bir araştırmada, uygulamadan mahrum kaldıkları için gücenen kişiler, misilleme yaparak kendi verimliliklerini düşürebilir ve deneysel grubu gerçekte olduğundan daha iyi gösterebilir. İkinci kirlenme etkisi ise *rekabet* olabilir. Kontrol grubundaki kişiler performanslarını düşürerek tepki göstermek yerine, çok çalışmak için motive olabilir ve performansları, uygulama grubundaki kişilerin performansından farklı çıkmaz. Son olarak, kirlenme *uygulamanın yayılmasına* yol açar. Eğer gruplar arasında uygulama bilgisine dair iletişim olursa (örneğin, öğle yemeği aralarında), kontrol grubundaki kişiler uygulamayı alanları taklit etmek için bu bilgiyi kendilerine uygulayabilirler. Örneğin, hastaların memnuniyetini arttırıp arttırmadığını test etmek için bir hastanede yeni bir düzenleme işlemi test edilebilir. Eğer uygulama yayılırsa, kontrol grubundaki personel bu işlem hakkında bilgi edinip kendi hastalarını düzenlemek üzere uygulayabilir. Bu durum deney ve kontrol grupları arasındaki hasta memnuniyeti farkını azaltır ve deneyin iç geçerliğine yönelik bir tehdit oluşturur.

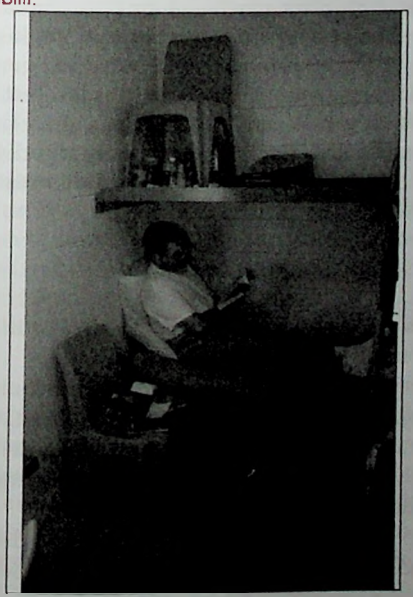
Gerçek deneyler, deneycinin farkında olmadan deney sonucunu etkilediği zaman oluşan *deneyci beklentisi etkisinden* de etkilenir. Gözlemci yanlılığı ise araştırmacıların yanlılıklarının ve beklentilerinin gözlemde, belirlemede, kayıtta ve davranışı yorumlamada sistemli hatalara yol açtığı zaman oluşur (Gözlemci ve deneyci yanlılığını çift körlemesine işlemi kullanarak kontrol etmenin çeşitli yolları 4'üncü ve 6'ncı bölümlerde özetlenmiştir).

Yenilik etkileri, deneysel bir uygulama gibi bir yenilik getirildiği zaman ortaya çıkabilir (Shadish ve ark., 2002). Örneğin, eğer bir iş yerinde ufak bir yenilik ya da değişim olursa, personel bu yenilik sunulduğu zaman heyecanlı ve enerjik hâle gelir. Uygulamanın kendisinden ziyade, personelin bu heyecanı, bu uygulamanın "başarı"sına neden olabilir. Yenilik etkisinin tam tersi, *bozulma etkisi* şeklinde ortaya çıkabilir; bu etkide bir yenilik, belki yeni bir çalışma yöntemi, personelin çalışmalarını o derece etkiler ki kendi olağan verimliliklerini bile gerçekleştiremezler.

Yenilik etkilerinin özel bir şekli, *Hawthorne etkisi* olarak bilinir. Bu etki, katılımcıların kendilerini "seçilmiş özel insanlar" olarak algılamalarından kaynaklanır. Bu etki, 1924 ve 1932 seneleri arasında, Şikago yakınlarındaki Cicero kasabasında kurulu Western Electric Şirketi'nin Hawthorne tesisinde yapılan araştırmadan sonra adlandırıldı (Roethlisberger, 1977). Verimlilik ve çalışma yeri koşulları arasındaki ilişkiyi incelemek için çalışmalar yürütüldü. Bir deneyde, tesisteki aydınlatma miktarı çeşitlendirildi ve personelin performansı incelendi. Sonuçlar, çalışma boyunca *hem deney hem de kontrol gruplarının* verimliliklerinin arttığını gösterdi. Bu etkiden sorumlu olan faktörlerin kesin olarak neler olduğu konusunda bir anlaşmazlık olmakla birlikte (örneğin, Parsons, 1974), Hawthorne etkisi genel olarak katılımcıların, kendileriyle ilgilenildiğini fark etmelerinden kaynaklanan bir davranışı gösterir.

Hawthorne etkisine örnek olarak tutukluların, hücre koşullarındaki değişiklik ile hapis hayatına karşı olan tutumları arasındaki ilişkiyi inceleyen bir çalışmaya katılmak için seçildiklerini düşünün (bk. Şekil 10.2). Eğer tutukluların tutumlarında olumlu değişiklikler elde edilirse, bu sonuç gerçekten hücre

ŞEKİL 10.2 Hapishane hayatının koşullarını iyileştirme yöntemlerini araştıran bir çalışma, Howthorne etkisinin konusu olabilir.



koşullarında yapılan değişikliklere bağlı olabilir ama tutukluların hapishane yönetiminin onlara gösterdikleri ilgi nedeniyle morallerinin yükselmesine de bağlı olabilir. Doğal ortamlarda çalışan araştırmacılar, katılımcıların davranışlarındaki değişikliklerin kısmen başkalarının onlara gösterdikleri ilgiye bağlı olabileceğinin farkında olmalıdır. Dolayısıyla, Hawthorne etkisinin önceki bölümlerde (özellikle 4'üncü Bölüm) bahsettiğimiz belirli bir tepkiselliği (örneğin, birinin gözlendiğine dair farkındalığı) temsil ettiğini görebilirsiniz.

İç geçerliğe olan tehditlerden doğan problemlere ek olarak, gerçek deneyler, *dış geçerliğe yönelik tehditler* tarafından da zayıflatılabilir. Dış geçerlik, bir çalışmanın sonuçlarının, çalışmada kullanılan insanların ve ortamların ötesine genellenebilme yeteneğidir. Kişi, ortam ve zaman örneklemleri, genelleme yapmak istediğimiz evreni temsil ettiği zaman dış geçerlik güçlenir. Temsil gücüne ulaşmanın en iyi yolu seçkisiz atama olmakla birlikte, araştırmacılar seçkisiz atamayı deneylerinde nadiren kullanırlar (Shadish ve ark., 2002). Bu nedenle, araştırmacılar bağımsız değişkenin etkisinin sadece belli örneklemdeki kişilere, ortamlara ya da zamana bağlı olabileceğinin farkında olmalılar. Şehir içindeki bir okulda kış döneminde gönüllü deney ve kontrol grupları arasında bir fark bulunduğunu varsayın. Araştırmanın dış geçerliği sorgulanabilir. Eğer gönüllü olmayan kişiler bahar döneminde bir banliyö okulunda test edilseydi aynı sonuçlar gözlenir miydi?

Dış geçerliği test etmenin en iyi yolu tekrar (replication) çalışmasıdır. Bu nedenle, dış geçerliğe ait sorular en iyi, deneyi farklı tip katılımcı, ortam ve uygulama yöntemleriyle ve farklı zamanlarda tekrar ederek cevaplanır. Zaman zaman "kısmî tekrar çalışmaları" deneye dâhil edilebilir. Sonuçların farklı grup insanlara genellenebilir olup olmadığını anlamak için birden fazla grup deneye dâhil edilebilir. Örneğin, bilişsel-davranışsal tedavi, bu uygulamanın etkisinin farklı gruplara genellenebilir olup olmadığına karar vermek için bir grup depresyon ve bir grup kaygı tanısı almış kişilerle test edilebilir. Benzer şekilde, uygulama etkisinin ortamlar arasında genellenebilir olup olmadığını görmek için, uygulama birden fazla klinikte test edilebilir. Sonuçların, terapistler arasında genellenebilir olup olmadığına karar vermek için birden çok terapist işe alınabilir. Gördüğünüz gibi dış geçerliğe dair sorular sayısız olabiliyor, ve bu, tekrar araştırmasını psikolojide araştırma yöntemlerinin önemli bir amacı yapıyor.

YARI DENEYSEL DESENLER

- Yarı deneysel desenler, gerçek deneylerin yürütülmesi mümkün olmadığı zaman önemli bir alternatiftir.
- Yarı deneysel desenlerde, gerçek deneylerde bulunan kontrol sağlanamaz; en önemlisi de seçkisiz atama yöntemi kullanılamaz.
- Araştırmacılar gerçek deneyden ziyade, yarı deneysel bir desende araştırma yaparken iç geçerliğe yönelik tehditleri yok etmek için ek bir kanıt aramalıdır.
- Tek grup ön test-son test deseni, ön-deneysel desen ya da kötü bir deney olarak adlandırılır çünkü iç geçerliği çok düşüktür.

Temel Kavram

Sözlüğe bakarsanız, *yarı* ön ekinin bir tanımının da "benzer" olduğunu göreceksiniz. Yarı deneysel desenler, gerçek deneylerle *benzer* işlemler içerir. Genel anlamda konuşursak, **yarı deneysel desenler** bir çeşit uygulama ya da müdahale ve karşılaştırmayı kapsar, fakat gerçek deneylerde bulunan kontrol derecesinden yoksundurlar. Nasıl ki seçkisiz atama gerçek bir deneyin ayırıcı özelliği ise, *seçkisiz atamanın olmayışı* da yarı deneysel desenin ayırıcı bir özelliğidir.

Araştırmacılar bir uygulamanın etkili olup olmadığını bilmek istedikleri fakat gerçek bir deneyin mümkün olmadığı durumlarda yarı deneysel desenlere başvururlar. Elde edilen bilginin az olması, hiç olmamasından daha iyidir ama yarı deneysel desenden sağlanan bir bilgiye güvenebilir miyiz? Bu soruyu cevaplamak için araştırmacılar dikkatlerini bu bölümde daha önce bahsedilen iç geçerliğe yönelik olası tehditlere verdi. Bir araştırmacı, elde ettiği bulgular için yapılabilecek olası farklı açıklamaları (yani tehditleri), yarı deneysel desenlerden gelen mevcut kanıtları inceleyerek ve durumun mantıksal analizini yaparak ortadan kaldırmanın yollarını arar. Örneğin, araştırmacı, çalışma boyunca katılımcıların ölçeklere verdikleri cevapları açıklayacak olası tarihî olaylar için durumu analiz edebilir. Tarih tehdidinin olasılık dışı olduğunu gösterebildiği zaman, araştırma sonuçlarının iç geçerliği için ortaya daha güçlü bir iddia koyabilir. Dolayısıyla, araştırmacılar kendi yarı deneysel desenlerinin eksik taraflarını gördükleri zaman dedektif gibi davranırlar ve bunların üstesinden gelmek için bulgu ararlar.

Bazı durumlarda, durumun mantıksal analizi ya da ek veri analizi yapılarak iç geçerliğe yönelik tehdidin giderilip giderilemeyeceğine karar vermek için yeterli delil yoktur. Bu durumda araştırmacılar kabul etmelidirler ki yarı deneysel desen, etkisi araştırılan faktörün etkiliği hakkında kesin bir delil sunmaz ve çalışmanın sonuçları için alternatif açıklamalar olabilir. *Yarı deneysel desen işlemlerinin sonuçlarından kaynaklanan yorumlama ile ilgili problemler yüzünden araştırmacılar, araştırma koşullarını gerçek deney koşullarına yaklaştırmak için her türlü gayreti göstermelidir.*

Doğal ortamlarda deney yürütürken araştırmacıların karşılaştığı belki de en ciddi sınırlılık, katılımcıları koşullara sıklıkla seçkisiz olarak atayamamalarıdır. Bu durum, bütün olarak bir grubun uygulama için diğerlerinden ayrı tutulduğu ve idarî kararların veya pratik hususların, kişilerin seçkisiz atanmasını engellediği zaman ortaya çıkabilir. Örneğin, bir okuldaki veya sınıftaki öğrenciler ya da bir tesisteki işçiler uygulama ya da müdahaleyi alacak olan toplu grubu temsil edebilir. Eğer bir grubun davranışının uygulamadan hem önce hem de sonra ölçüldüğünü varsayarsak, böyle bir "deney" şöyle betimlenir:

$$O_1 \quad X \quad O_2$$

Bu ifadedeki O_1, grubun ilk gözlemini ya da ön testi, X uygulamayı, O_2 ise ikinci gözlemi ya da son testi temsil etmektedir.

Bu *tek grup ön test-son test deseni*, ön deneysel deseni veya kötü bir deneyi temsil eder. Ön test ve son test skorları arasında elde edilen herhangi bir fark uygulamaya *ya da* tarihe, olgunlaşmaya, sınama etkisine ve araç düzeni dâhil iç geçerliğe yönelik olabilecek bir tehdide bağlı olabilir (deneyci beklentisi ve yenilik etkilerinin yanı sıra). Araştırmacılar böyle kötü

bir deneyde uygulamanın etkililiği hakkında herhangi sonuca varamazlar. Neyse ki ön deneysel desenlerden çok daha iyisini yapabilen yarı deneysel desenler vardır.

Denk Olmayan Kontrol Grup Deseni

- Denk olmayan kontrol grup deseninde, uygulama grubu ile karşılaştırma grubu ön test ve son test ölçümleri kullanılarak karşılaştırılır.
- Eğer bu iki grubun uygulama öncesi ön test puanları benzer, uygulama sonrası elde edilen son test puanları farklı ise, araştırmacılar uygulamanın etkili olduğunu daha emin bir şekilde iddia edebilirler.
- Denk olmayan grup deseninde tarih, olgunlaşma, sınama ve araç düzenine bağlı iç geçerliğe yönelik oluşabilecek tehditler kontrol edilebilir.

Temel Kavram

Eğer iki koşul karşılanabilirse, tek grup ön test-son test deseni yüksek bir iç geçerliği olan bir yarı deneysel desene dönüştürülebilir: (1) Karşılaştırma grubuna "benzer" bir uygulama grubu olduğu zaman ve (2) hem uygulama hem de karşılaştırma grubundaki bireylerden ön test ve son test ölçümleri elde etme fırsatı olduğu zaman. Campbell ve Stanley (1966) bu iki koşulu karşılayan yarı deneysel işleme, **denk olmayan kontrol grup deseni** demiştir. Karşılaştırma grubu seçkisiz yöntemle atanmadığı için, uygulama ve karşılaştırma gruplarındaki kişilerin bütün önemli özellikler bakımından eşit olduğunu varsayamayız (bir başka deyişle seçime yönelik tehditler ortaya çıkar). Bu nedenle, ön testin, bağımlı değişken ölçümlerinin ne kadar benzer

İLERİ ALIŞTIRMA

Bu egzersizde sizin, aşağıda kısaca betimlenmiş olan tek grup ön test-son test deseninde iç geçerliğe yönelik olası tehditleri düşünmenizi istiyoruz.

Yeni bir depresyon tedavisinin etkisini araştıran bir psikolog, tedavi olmak isteyen 20 depresyon hastasını araştırmaya aldı. Çalışmanın başında bütün katılımcılardan depresyon semptomları hakkında bir anket doldurmalarını istedi. Tüm örneklem için ortalama depresyon puanı 42.0'dır (olası en yüksek puan 63.0'dır) ki bu puan şiddetli semptomları işaret eder. (Depresyonda olmayan kişilerin bu ölçekteki puanları genellikle 0 ile 10 arasındadır.). Sonraki 16 hafta boyunca psikolog katılımcıları yeni yöntemle tedavi etti. Tedavi sürecinin sonunda katılımcılar depresyon ölçeğini tekrar doldurdu. Son test puanlarının ortalaması 12,0'dır. Bu puan katılımcıların depresif semptomlarının belirgin şekilde azaldığını ve hafif depresyonu işaret eder. Psikolog tedavinin etkili olduğu sonucuna vardı; yani, tedavi depresif semptomları geriletti.

Bu psikoloğun yaptığı gibi bir neden-sonuç ifadesi kullanmak, tek grup ön test-son test deseni kullanırken aslında imkânsızdır. Bunun neden böyle olduğunu anlamak için bu çalışmadaki iç geçerliğe olası tehditleri düşünmenizi istiyoruz.

1. Tarih etkisi bu çalışmanın iç geçerliğini nasıl tehdit edebilir?
2. Olgunlaşmanın bu çalışmada nasıl bir rol oynayabileceğini açıklayın.
3. Test etkisi ve araç düzeni bu çalışma için olası tehditler olabilir mi?
4. İstatistiksel regresyonun, bu sonuçların yorumunu nasıl etkileyebileceğini açıklayın.

olduğunu ölçmek için her iki gruba da verilmesi esastır. Denk olmayan kontrol grup deseni şöyle özetlenebilir:

$$\begin{array}{ccc} O_1 & X & O_2 \\ \text{------} & & \\ O_1 & & O_2 \end{array}$$

Buradaki kesik çizgiler, katılımcıların müdahale ve kontrol gruplarına seçkisiz olarak atanmadığını gösteriyor.

Araştırmacılar tarih etkisine, olgunlaşmaya, sınama etkisine, araç düzenine ve regresyona bağlı olarak iç geçerliğe yönelik oluşan tehditleri bir kontrol grubu ekleyerek kontrol edebilir. Deneysel bir desenin mantığını kısaca değerlendirirsek, neden böyle olduğunu daha iyi anlayabiliriz. Deneye iki benzer grupla başlamak isteriz; sonra ise bu gruplardan biri müdahaleye maruz kalırken, diğeri kalmaz. Eğer bu iki grubun son test puanları müdahale sonrası farklılaşırsa, bu farklılığın sebebinin müdahale olduğunu gösterebilmek için, önce alternatif açıklamaları elemek zorundayız. Eğer gruplar tamamen denk ise ve her iki grup da benzer deneyimler yaşamışsa (müdahale hariç) tarih etkisi, olgunlaşma, sınama etkisi, araç düzeni ve regresyon etkilerinin her iki grupta da eşit olduğunu varsayabiliriz. Dolayısıyla, her iki grubun bu süreçte aynı oranda değiştiğini (olgunlaşma), çoklu değerlendirmelerden benzer şekilde etkilendiklerini, ya da benzer dış olaylara maruz kaldıklarını varsayabiliriz (tarih etkisi). Eğer bu etkiler her iki grupta da *benzer* şekilde deneyimlenirse, son test puanlarındaki *farklılığı* açıklamak için muhtemelen kullanılamazlar. Bu nedenle, artık iç geçerliğe yönelik tehdit oluşturmazlar. Bu nedenle araştırmacılar, denk bir grup ekleyerek, nedensel çıkarımlar yapmak için önemli bir avantaj kazanırlar. Bununla birlikte bu nedensel çıkarımlar çalışmanın başında denk gruplar oluşturmaya ve bu grupların müdahale dışında benzer deneyimler yaşadığından emin olmaya dayanır. Çünkü, göreceğimiz gibi, bunu uygulamada fark etmek zordur; seçimle gelen etkilere bağlı iç geçerliğe yönelik tehditler, bu desende ortadan kaldırılmaz

Denk olmayan kontrol grup deseni kullanan bir çalışma örneği için, araştırma yöntemleri dersini bir deneysel "müdahale" olarak almanın etkisini ele alın (VanderStoep & Shaughnessy, 1997). Araştırma yöntemleri dersine giren öğrencilerle gelişim psikolojisi dersine giren öğrencilerin günlük olaylar hakkında kritik düşünme becerisi karşılaştırıldı. Öğrenciler dersin başında yapılan ön testte muhakeme becerileri bakımında benzer puanlar aldılar, fakat dönem sonunda araştırma yöntemi alan öğrenciler kontrol grubundaki öğrencilerden daha büyük gelişme gösterdiler. Bu bir parça cesaret verici örneği aklımızda tutarak, denk olmayan kontrol grup deseni kullanan bir başka çalışmayı detaylı olarak inceleyelim. Bu, bize yarı-deneysel işlemlerin hem güçlü yanlarını hem de sınırlılıklarını gözden geçirme imkânı verecektir.

Denk Olmayan Kontrol Grup Deseni: Langer ve Rodin'in Çalışması

- Yarı deneysel desenler genellikle birçok unsuru olan bir müdahalenin tüm etkililiğini ölçerler; daha sonra yapılacak takip (follow-up) çalışması hangi unsurların müdahalenin etkisine ulaşmak için kritik olduğuna karar verebilir.

Langer ve Rodin (1976) huzurevlerinde yaşlı insanların fiziksel ihtiyaçları büyük ölçüde karşılandığı hâlde bu tip kurumların orada yaşayanların kendi kişisel kararlarını verebilmeleri için çok az şans tanıdığını gözlemledi (bk. Şekil 10.3). Çok basit seçimleri bile yapma imkânının az olmasının bazen yaşlılarda psikolojik hatta fiziksel güçten düşmeye neden olabileceğine dair bir hipotez kurdular. Hipotezlerini test etmek için, Connecticut Huzurevi'nde denk olmayan kontrol grup deseni kullanarak yarı deneysel bir deney yürüttüler.

Deneydeki bağımsız değişken, huzurevinde kalan iki gruba verilen farklı tipteki sorumluluktu. Sorumluluk verilen gruba, odalarının nasıl olacağı, ziyaretleri, film seçimleri ve benzeri konularda kararlar vermeleri gerektiği bildirildi. Bu kişilere aynı zamanda hediye olarak birer küçük bitki verildi (bunu kabul edip etmeme kararı da onlara aitti) ve eğer dilerlerse ona bakabilecekleri söylendi. Karşılaştırma grubu ise bir araya getirilerek personelin onlar için karar verme sorumluluğu bildirildi. Bu gruptaki kişilere onlar isteseler de istemeseler de birer bitki verildi ve personelin bakacağı söylendi.

Huzurevi sakinleri bu iki gruptan birine seçkisiz olarak atanmadı, bunun yerine iki tip sorumluluk derecesi kendilerine kaldıkları iki farklı kata göre verildi. Langer ve Rodin (1976) katların "sakinlerin fiziksel ve psikolojik benzerlikleri ve önceki sosyo-ekonomik statüleri açısından" seçildiğini bildirmiştir.

ŞEKİL 10.3 Langer ve Rodin (1976) iki farklı sorumluluk tipinin, huzurevinde yaşayanların davranışları üzerindeki etkisini araştırmak için denk olmayan kontrol grubu deseni kullandı. "Gerçek bir deney" yürütülmediği için araştırmacılar iç geçerliğe herhangi bir tehdit olup olmadığına karar vermek için araştırmanın özelliklerini incelediler.



(s. 193). Anketler, çalışmaya katılan sakinlere sorumluluk ile ilgili aldıkları yönergelerden bir hafta önce (ön test) ve yönergeden üç hafta sonra (son test) verildi. Anketteki maddeler kişilerin kontrol hissi, mutluluğu ve aktivite dereceleri hakkındaydı. Ayrıca, her bir kattaki görevliler çalışmaya katılan kişilerin uyanıklık, sosyallik ve aktiflik düzeylerini müdahaleden önce ve sonra derecelendirdiler. Son olarak araştırmacılar sosyal ilgiyi ölçmek için akıllıca bir son test ölçümü yapıp bir yarışma düzenlediler. Çalışmaya katılan huzurevi sakinlerinden büyük bir kavanozda bulunan şekerleme sayısını tahmin etmelerini istediler. Yarışmaya katılmak isteyen sakinler bir kâğıda isimlerini ve tahmin ettikleri sayıyı yazdılar.

Sorumluluk verilen gruptaki kişiler, karşılaştırma grubuna kıyasla, genel olarak daha mutlu, daha hareketli, daha dikkatli idiler ve sorumluluk verilen gruptan 10 kişi yarışmaya katıldığı hâlde, karşılaştırma grubundan sadece 1 kişi katıldı! Langer ve Rodin araştırmalarının sonuçlarına dayanarak yaşlanmanın bazı olumsuz etkilerinin kişilere kendi kararlarını verme olanağı sağlanarak azaltılabileceğini ya da tersine çevrilebileceği sonucuna vardı.

Denk olmayan kontrol grup deseninin belirli sınırlılıklarına geçmeden önce, Langer ve Rodin'in çalışmasının bir başka yönü bahsetmeye değer. Araştırmalarındaki müdahale birden fazla unsur içerdi: Huzurevi sakinleri çeşitli kararlar verme konusunda cesaretlendirildi (örneğin; filmler, odalar ve aktiviteler) ve bakmaları için bitki verildi. Bununla birlikte deney, bütün bu müdahaleyi ayrı ayrı unsurlarıyla değil, tek bir "paket" olarak değerlendirdi. Müdahaleyi oluşturan hangi unsurların gerçekten etkili olduğunu veya hepsine ihtiyaç olup olmadığını ya da unsurlardan birinin diğerlerinden daha önemli olup olmadığını bilmiyoruz.

Doğal düzeneklerdeki müdahalelerin genelde birden fazla unsuru olur. Örneğin travma sonrası stres bozukluğu (TSSB) alanındaki bir müdahale "paketi" bilgilendirici broşürler, sağlık hizmeti uzmanları ile bireysel terapi, destek grubu toplantıları ve yogayı kapsayabilir. Birçok çalışmanın temel amacı bu unsurların toplam etkisini ölçmektir. Eğer müdahale paketinin etkili olduğu gösterilirse, sonraki araştırma o paketin özellikle hangi unsurun etkili olduğunu belirleyebilir. Kritik unsurları belirlemek, olumlu sonuca katkı sağlamayan unsurların çıkarılmasıyla, maddî destek sağlayan kişilere müdahale masrafını azaltma imkânı verir. Bununla birlikte, pratik değerlendirmelere ek olarak, araştırmacılar davranış hakkındaki bir teoriye dayanan bir müdahaleye ilişkin yordamaları test ettiklerinde, psikolojik teorilerin gelişimi de bir rol oynar. Bir müdahalenin faydalı bir etkisini gösteren bir çalışma duyduğunuzda ya da okuduğunuzda, müdahalenin birden çok unsur içerip içermediğine dikkatlice bakın ve olumlu müdahale etkisi üretmek için gerekli unsurları ortaya çıkaracak ek ne gibi araştırmalar yapılabileceğini düşünün.

Denk Olmayan Kontrol Grup Deseninde İç Geçerliğe Tehditler

Yarı deneysel desenlerdeki sonuçları yorumlamak için, araştırmacılar iç geçerliğe herhangi bir tehdit olup olmadığını inceler.

- Denk olmayan kontrol grubu deseni kullanılırken göz önünde bulundurulması gereken iç geçerliğe tehditler şunlardır: Seçimin ilave etkileri, gözlemci yanlılığı, kirlenme ve yenilik etkisi.
- Gruplar ön testte denk olabilmelerine rağmen, bu durum grupların çalışma sonucuyla ilişkili her hususta denk olabileceklerini garanti etmez.

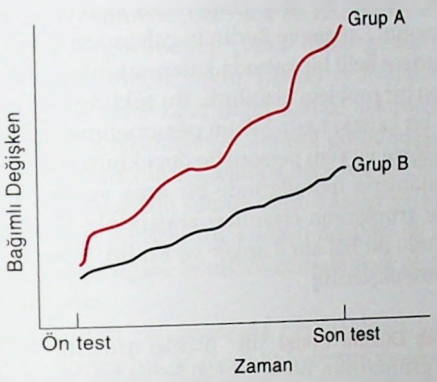
Cook ve Campbell'e (1979) göre denk olmayan kontrol grup deseni, iç geçerliğe yönelik potansiyel tehditlerin başlıcalarını kontrol eder ama 1) seçim ve olgunlaşmanın, 2) seçim ve geçmişin, 3) seçim ve araç kullanmanın ve 4) diferansiyel istatistiksel regresyonun (yani, ortalamaya gerilemenin bir grupta gözlenmesi ama diğerinde gözlenmemesinin) etkisine bağlı olan tehditleri kontrol etmez. Geçerliğe yönelik sorun yaratacak bu her bir tehdidin Langer ve Rodin'in çalışmasında ne gibi sorunlar yaratabileceğini inceleyeceğiz. Sonra, Langer ve Rodin'in çalışmalarının iç geçerliğine yönelik olası tehditleri çürütmek için nasıl akla uygun iddialar ve deneysel kanıtlar önerdiklerini açıklayacağız. Aynı zamanda deneyci yanlılığı ve kirlenmenin nasıl kontrol edildiğini inceleyeceğiz. Son olarak denk olmayan kontrol grup deseninin doğasında olan dış geçerliği sağlamanın zorlukları üstünde kısaca yorum yapacağız.

Langer ve Rodin'in çalışmasında önemli bir ön bulgu iki gruptaki huzurevi sakinlerinin ön test ölçümlerinin anlamlı olarak farklı çıkmamasıydı. Bu iki grup arasında müdahaleden önce fark bulmak pek şaşırtıcı olmazdı çünkü kişiler koşullara seçkisiz olarak atanmamıştı. Bununla birlikte, ön test sonuçları gruplar arasında herhangi bir farklılık olduğunu göstermese de bu grupların "denk" olduğunu varsayamayız (Campbell & Stanley, 1966). Grupların denk olduğu sonucunu neden çıkaramayacağımızı aşağıda açıklayacağız.

Seçim-Olgunlaşma Etkisi Olgunlaşma etkisi bir grupta meydana gelen doğal değişme oranı demektir. Seçim ve olgunlaşmanın ilave etkisi, bir gruptaki kişiler diğer gruptakilerden farklı bir oranda değiştiği zaman ortaya çıkar (Shadish ve ark., 2002). Örneğin bir gruptaki bireyler, diğer gruptaki bireylerden daha fazla deneyim edinebilir, yorulabilir, sıkılabilir ya da daha hızlı bir oranda sağlıksız hâle gelebilir. Kişilerin araştırmada bir müdahale istedikleri ve karşılaştırma grubunun müdahale grubundan farklı bir evrenden seçildiği hâllerde, seçme-olgunlaşma etkisi iç geçerliğe yönelik daha olası bir tehdittir (Campbell & Stanley, 1966).

Müdahale ve karşılaştırma grubunun ön test skorlarının yaklaşık aynı olduğu zamanlarda bile birbirlerine denk olduğu sonucuna varamamamızın bir nedeni, seçim-olgunlaşma etkisi olasılığıdır. İki grup için doğal değişme oranı farklılaşabilir (iki grubun aynı evrenden olmadığı veya denk olmadığı anlamında), ama ön test ölçümü iki grup yaklaşık aynı noktada olduğu zaman da verilebilirdi. Bir araştırmacı muhtemelen, grubun olgunlaşma oranını sadece tek bir ön test puanına dayanarak bilemez. Bu problem Şekil 10.4'te gösterilmiştir. Grup A'nın gerçek değişme oranının Grup B'den daha büyük olduğunu varsayın (farklı eğimli çizgilerle gösterilmiştir), fakat ön test verildiği zaman iki grup hemen hemen aynı noktadadır. Araştırmacı kolayca bu iki grubun denk

ŞEKİL 10.4 Müdahale yokluğunda iki grup (A ve B) için muhtemel farklı büyüme oranları.



olduğu sonucuna varabilir, fakat aslında değişme oranları farklıdır. Ayrıca, bağımlı değişken için son test sonuçları iki grubun basitçe farklı olduğunu işaret eder çünkü olgunlaşma oranları farklıdır. Sadece ön test ve son test gibi iki ölçümü kullanan bu durumdaki bir araştırmacı, yanlış şekilde bu iki grup arasında çıkan son testteki farklılığı müdahalenin yarattığı sonucuna varabilir.

Aynı zamanda şu hususu da akılda tutmak önemlidir: İki grup, çalışmanın başında yapılan ön testte ya da başka herhangi bir grup testinde denk olsalar bile, öngörülemeyen başka şekillerde farklılaşabilir. Bu farklılıklar yarı deneysel deneylerde kişilerin davranışlarıyla ilişkili olabilir ve araştırma bulguları için alternatif açıklama seçenekleri (diğer bir deyişle iç geçerliğe yönelik tehdit) olabilir.

Langer ve Rodin'in çalışmasında seçim-olgunlaşma etkisinden şüphelenmemiz için neden olabilir mi? Müdahale katında yaşayan huzurevi sakinlerinin, karşılaştırma katında yaşayanlardan doğal şekilde daha farklı bir oranda değişmiş olmaları mümkün müdür? Bulguların bir çoğu bunun pek mümkün olmadığını düşündürüyor. İlk olarak, sakinler odalarına atanırken kullanılan prosedür seçkisizdi. Eğer sakinler huzurevinin katlarına farklı düzeyde duydukları ihtiyaca göre yerleştirilmiş olsalardı, bu kesinlikle bir sorun olurdu (yani, muhtemelen farklı yaşlanma oranlarını gösteren), fakat bu çalışmada durum böyle değildi. İkinci olarak, Langer ve Rodin iki farklı katı müdahale veya karşılaştırma grubuna seçkisiz atayarak çalışmalarını gerçek bir deneye yaklaştırmış oldu. Sorumluluk verilen grup özellikle böyle bir müdahale istemedi ve çeşitli ön test ölçümlerinde demografik bilgilerde ortaya çıkan gruplar arası benzerliğe dayanarak Langer ve Rodin grupların aynı popülasyondan olduğu sonucuna vardı. Dolayısıyla, bu iki grubun doğal değişme oranlarının farklı olduğuna dair çok az bulgu var; bu da araştırmacıları, seçim-olgunlaşma faktörünü iç geçerliğe tehdit olarak görmemeye yöneltir.

Seçim-Tarih Etkisi Denk olmayan kontrol grup deseninde kontrol edilmeyen bir diğer iç geçerliğe yönelik tehdit, seçim ve tarihin ilave etkisidir. Cook ve

Campbell (1979) bu problemi yerel tarih etkileri olarak adlandırdı. Bu problem, müdahale dışında bir olay gruplardan sadece birisini etkilediği zaman ortaya çıkar. Örneğin, Langer ve Rodin'in çalışmasında eğer geçmişteki bir olay huzurevinin sadece belli bir katında kalan sakinlerin mutluluk ve canlılıklarını etkileseydi, bu bir problem olabilirdi. Bu şekilde bir çok olasılık hayal edebilirsiniz. Örneğin, bir kattaki hasta bakım personelinin değişmesi gibi. Böyle bir değişikliğin niteliğine ve yeni personelle öncekinin davranışları arasındaki farka bağlı olarak, sakinlerin morallerinde bir artış veya düşüş olabilir. Karşılaştırma ve müdahale gruplarının ortamları arasındaki farklılık arttıkça, yerel tarih etkisi daha sorunlu bir hâl alır. Langer ve Rodin, yerel tarih etkisi meselesine özel bir önem vermemişlerdir.

Seçim-Araç Düzeni Etkisi Bir ölçme aracındaki değişmelerin keşfedilmesi olasılığı, gruplardan birisinde diğerinden daha fazla olduğu zaman, seçim ve araç kombinasyonuna bağlı bir tehdit ortaya çıkar. Örneğin, taban ve tavan etkileri ön test ve son test arasındaki davranış değişikliğini belirmeyi zorlaştırabilir. Eğer bu durum bir grupta diğerine nazaran daha sorunluysa, seçim-araç düzeni etkisi vardır. Shadish ve arkadaşları (2002), iç geçerliğe yönelik bu tehdidin, denk olmayan gruplar ile puanları uç değerde olan grup için problem olması olasılığının daha büyük olduğuna işaret etti. Langer ve Rodin'in grupları ön testte farklılaşmadığı ve gruplar kullanılan ölçeklerde taban ya da tavan etkisi göstermediği için böyle bir tehdidin olması olası görünmüyor.

Diferansiyel İstatistiksel Regresyon Denk olmayan kontrol grubu deseninde iç geçerliğe yönelik son tehdit diferansiyel istatistiksel regresyondur (Shadish ve ark., 2002). Daha önce de ifade ettiğimiz gibi, ortalamaya doğru gerileme (regresyon), kişilerin aldıkları uç puanlara göre seçildikleri durumlarda beklenir (örneğin en zayıf okuyucular, en düşük verimlilikle çalışan işçiler, en ciddi problemleri olan hastalar gibi). *Diferansiyel regresyon*, bir gruptaki gerileme olasılığı diğerinden daha yüksek olduğu zaman oluşabilir. Örneğin, problemleri en ciddi olan hastaların müdahale grubuna yerleştirildiği denk olmayan bir kontrol grup deseni düşünün. Bu grup için bile gerileme olasıdır. Eğer gerileme, müdahale grubunda, karşılaştırma grubuna nazaran daha olası ise, ön test ve son test arasındaki fark yanlışlıkla müdahale etkisi olarak yorumlanabilir. Çünkü Langer ve Rodin'in çalışmasındaki gruplar aynı evrenden geldi ve gruplardan birinin ön test puanının diğerinden daha uç olduğuna dair herhangi bir bulgu yok. Bu çalışmada, iç geçerliğe yönelik diferansiyel istatistiksel regresyona bağlı bir tehdit olasılığı yoktur.

Beklenti Etkileri, Kirlenme ve Yenilik Etkileri Langer ve Rodin'in çalışması gerçek deneyleri bile etkileyebilecek iç geçerliğe yönelik ilave üç faktörden etkilenmiş olabilir: Beklenti etkileri, kirlenme ve yenilik etkileri. Eğer araştırmadaki gözlemciler (örneğin, çalışanlar) araştırma hipotezinin farkına varmışlarsa,

sorumluluk verildikten sonra huzurevi sakinlerinin farkında olmadan daha iyi olduklarını farz etmeleri mümkündür. Ama bu gözlemci yanlılığı veya beklenti etkisi kontrol edilmiş gibi görünüyor, çünkü hiçbir gözlemcinin araştırma hipotezine dair herhangi bir bilgisi yoktu. Langer ve Rodin aynı zamanda olası kirlenme etkilerini de ele aldı. Eğer diğer kattaki sakinlere kendi kararlarını verme fırsatı verildiğini öğrenmişlerse, kontrol grubundaki sakinlerin moralleri bozulmuş olabilir. Bu durumda aslında huzurevinin farklı katlarını farklı koşullar için kullanmak avantajlı olmuş oldu; Langer ve Rodin'in (1976) belirtiğine göre "Katlar arası iletişim söz konusu değildi" (s. 193). Bu yüzden en azından çalışmanın iç geçerliğine tehdit oluşturabilecek ölçekte kirlenme etkisi var gibi görünmemektedir.

Eğer Langer ve Rodin'in çalışmasında müdahalenin yapıldığı kattaki sakinler bu yenilikçi ve sorumluluk alabildikleri müdahale nedeniyle daha heyecanlı ve enerjik olsalardı, bu çalışmada yenilik etkileri olurdu. Dolayısıyla, sorumluluk duygusundan ziyade sakinlerin yaşadıkları bu heyecan müdahale etkisini açıklayabilir. Ayrıca, müdahale grubuna verilen bu özel dikkat, grupta bir Hawthorne etkisi yaratmış olabilir ve müdahale katında yaşayan sakinler kendilerini daha iyi hissetmiş olabilirler. Bu çalışmada Hawthorne etkisini ya da yenilik etkilerini tamamen göz ardı etmek oldukça zordur. Bununla birlikte araştırmacılara göre "İki gruba verilen dikkat miktarında hiçbir fark yoktu" (s.194). Aslında her iki grupla da sağlanan iletişimin vurgulandığı nokta, personelin her iki grupla da ilgilendiği ve onların "mutlu olmalarını" istedikleriydi. Dolayısıyla aksi herhangi bir bulgu olmadığı sürece, Langer ve Rodin'in gözlediği davranış değişikliklerinin, kontrol edilemeyen bir dış etkenden ziyade, bağımsız değişkenin etkisi olduğu sonucuna varabiliriz.

Belirli bir deneyde bağımsız değişkenin "işleyip işlemediğine" karar vermek için araştırmacılar, müdahalenin davranışı değiştirdiğinin lehinde ve aleyhinde olan bulguları sistemli bir şekilde toplayıp dikkatlice değerlendirmelidir. Bu süreçte iç geçerliğe yönelik tehdit listesini gözden geçirilir, bağımlı ve bağımsız değişken arasındaki bir ilişkiyi açıklayabilecek makul açıklamaları eleme kararı vermek için veriler ve durum incelenir. Araştırmacı ancak bütün tehditler elendikten sonra müdahalenin etkili olduğunu emin bir şekilde öne sürebilir.

Dış Geçerlik Meselesi

- Dış geçerlik meselesi de iç geçerlikte olduğu gibi, ciddi bir şekilde incelenmelidir.
- Araştırma bulgularında dış geçerliğin en iyi kanıtı farklı evrenlerle, farklı ortamlarda ve zamanlarda yapılan tekrar çalışmalarıdır.

İç geçerliğe yönelik yaptığımız sistemli analizin aynısını yarı-deneysel desenlerin dış geçerliklerine yönelik de yapmalıyız. Belirli sonuç örüntüleri sadece belli bir grup katılımcı, ortam, ve zaman için mi geçerlidir? Örneğin, Langer ve Rodin yaşlı insanların bakım şeklinde belli değişikliklerin yapıldığını

belirtmekle birlikte, sorumluluk veren müdahalenin bütün yaşlı huzurevi sakinleri için ve tüm huzurevleri için geçerli olup olmadığını sorgulayabiliriz. Langer ve Rodin'in (1976) seçtiği huzurevinin "Connecticut eyaletinin en iyi bakım üniteleri arasında" (s. 193) olarak tanımlanması, huzurevi sakinlerinin, olanaklarının ve personelinin diğer benzerlerinden farklı olabileceğini akla getiriyor. Örneğin, eğer bu huzurevi sakinleri buraya gelmeden önce diğer huzur evlerine giren sakinlere göre daha bağımsız idiyseler (muhtemelen, sosyo-ekonomik statü farklılığı nedeniyle), taşınma deneyiminden daha fazla etkilenmiş olabilirler. Neticede, sorumluluk almaları, başka huzurevlerine gelen sakinlere kıyasla onlar için daha önemli olmuş olabilir.

Araştırmacılar bulgularının genellenebilir olup olmadığını sorguladıkları zaman, dış geçerliği sağlamak için bulgularını farklı evren, ortam ve zamanlarda *tekrarlama çalışması* (replication) yapma fırsatı yaratırlar. Örneğin, sorumluluk veren müdahale, farklı huzurevlerinde ya da belki engelli kişilerle veyahut hastane dışı tedavi merkezlerinde yaşayan ergenlerle denenebilir. Sorumluluk veren müdahalenin uygulandığı bu yöntemde, müdahale belli bir ortam ve popülasyona adapte edilmek üzere değiştirilebilir. Gördüğünüz gibi bulgunun dış geçerliği üstünde düşündükçe yeni araştırma soruları ortaya çıktığını görüyorsunuz. Tek başına hiçbir çalışma, araştırma hipotezi ile ilgili bütün sorulara cevap veremez ve araştırma birikimli bir etkinliktir: Bulgular çok sayıda araştırmayla tekrarlandıkça hipotez ve teorilerimiz destek kazanır.

Kesintili Zaman Dizileri Deseni

- Basit kesintili zaman dizileri deseninde araştırmacılar hem müdahale öncesi hem de sonrasında bir dizi gözlem yaparlar.
- Zaman dizileri deseninde müdahalenin etkisini gösteren kanıt, müdahale yapıldığında ani bir değişiklik (düzensizlik) olduğu zaman oluşur.
- Basit kesintili zaman dizileri deseninde iç geçerliğe yönelik en büyük tehditler müdahale ile aynı zamanda ortaya çıkan tarih etkisi ve ölçekteki değişikliklerdir (araç düzeninde).

Temel Kavram

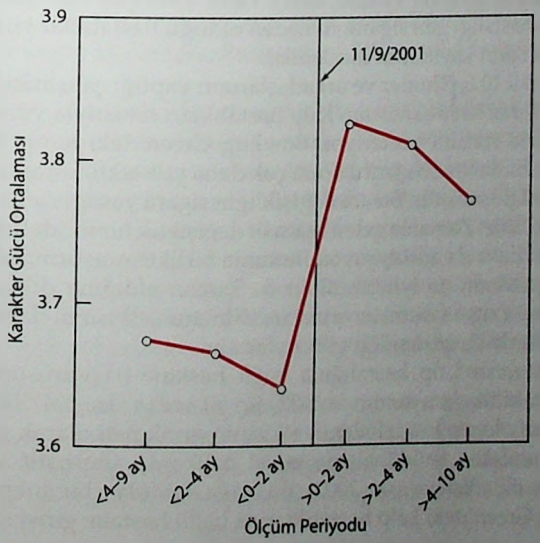
Araştırmacılar bağımlı değişkeni müdahale öncesi ve sonrası inceleyebildikleri zaman, ikinci bir yarı deneysel desen olarak **basit kesintili zaman dizileri deseni** mümkündür (Shadish ve ark., 2002). Gözlemler aynı kişiler üzerinde tekrar edilebilir (örneğin, tıbbî bir müdahalenin hastaların semptomları üzerindeki etkisi değerlendirilirken) ya da zaman dizileri, bir bağımlı değişkenin evrendeki gözlemlerini kapsayabilir. Örneğin, araştırmacılar trafik kazası sonucu ölüm dosyalarını, video oyunu satışlarını ya da kiliseye devam arşivlerini inceleyebilirler –liste sonsuzdur. Basit kesintili zaman dizileri çok çeşitli müdahale tiplerinin etkililiğini ölçmek için kullanılır; yeni bir ürün tanıtılırken, yeni sosyal bir reform başlarken ya da bir hastanede enfeksiyonu azaltmak için bir

program başlatırken. Bu desenin özü müdahale öncesi ve sonrasında yapılan düzenli ölçümdür. Basit kesintili zaman dizileri şu şekilde özetlenebilir.

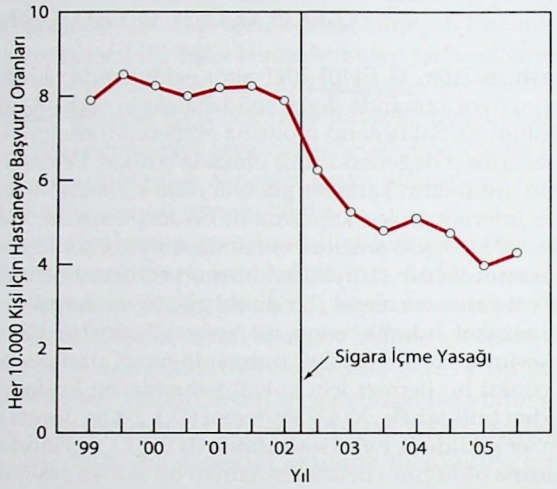
$$O_1\ O_2\ O_3\ O_4\ O_5\ X\ O_6\ O_7\ O_8\ O_9\ O_{10}$$

Araştırmacılar, 11 Eylül 2011 terör saldırısında olduğu gibi, kesintili zaman dizilerini aynı zamanda doğal müdahalelerin etkisini ölçmek için de kullanırlar (bk. Bölüm 4). Saldırıların olumsuz etkilerini (örneğin; korku, ön yargı gibi) bir çok araştırmacı değerlendirmiş olmakla birlikte, Peterson ve Seligman (2003) saldırıların, insanların karakter gücünü nasıl etkilediğini araştırdı. Devam etmekte olan bir internet projesi kapsamında bu araştırmacılar *Values in Action (VIA) Classification of Strengths* anketini tamamlayan yüzlerce kişinin arşiv bilgisini elde ettiler Araştırmacılar 11 Eylül saldırısının etkilerini betimlemek amacıyla, şu yedi gücün ortalamasını alarak, bir dinsel güç ölçüm formu hazırladılar: Minnettarlık, umut, nezaket, liderlik, sevgi, maneviyat (kader) ve takım çalışması. Çalışmaları basit kesintili zaman dizileri deseninin temel özelliklerini iyi bir şekilde gösteriyor çünkü bu itemler için anketi tamamlayan kişilerin verdikleri cevapları 11 Eylül'den (müdahale, X) 4-9 ay öncesi (O_1), 2-4 ay öncesi (O_2) ve 0-2 ay öncesi (O_3) ve benzer şekilde 11 Eylül sonrasında da (O_4,O_5,O_6) incelediler. (Bunun boylamsal bir çalışma olduğunu belirtelim, çünkü bu zaman aralıkları boyunca anketi farklı kişiler doldurdu). Şekil 10.5 her bir zaman aralığında, her dinsel güç için verilen cevapların ortalamalarını gösteriyor (daha yüksek puanlar, 5 dereceli ölçekte daha büyük gücü temsil ediyor). Grafikte açıkça görünen kesinti (ani değişiklik)

ŞEKİL 10.5 11 Eylül 2011 öncesi ve sonrası özbildirime dayanan karakter güçlerine (minnettarlık, umut, nezaket, liderlik, sevgi, maneviyat ve takım çalışmasının bileşimi) ait ortalama puanlar. Yüksek skorlar yüksek gücü temsil etmektedir (Ortalamalar, Peterson & Seligman, 2003, Şekil 1'den değiştirilerek alınmıştır.).



ŞEKİL 10.6 1999 ve 2005 yılları arasında Bowling Green'de her 10.000 kişi içinde kalp rahatsızlığından hastaneye yatanların 6 aylık ortalaması (Ocak-Haziran, Temmuz-Aralık) Veriler, Khuder ve arkadaşlarının (2007) bulgularından alınmıştır.



11 Eylül saldırılarının kişilerin öz bildirimine dayanan güçlerindeki artışa dair kanıt sağlıyor.

Zaman dizilerindeki kesinti, müdahalenin etkisine dair en büyük kanıt olmakla birlikte, araştırmacılar gözlemlerinde genellikle kademeli bir değişiklik gözler. Gözlemler aynı zamanda büyük çeşitlilik gösterebilir. Ohio'daki araştırmacılar, 2002'de iş yerlerinde ve umuma açık alanlarda sigara içilmesini yasaklayan Bowling Green'in temiz hava yasalarının etkilerini araştırdıkları zaman, durum böyleydi (Khuder ve ark., 2007). Ocak 1999'dan Haziran 2005'e kadar koroner kalp hastalığı gibi sigaranın neden olduğu hastalıklar yüzünden hastaneye başvuranların kayıtlarını incelediler.

Şekil 10.6, Khuder ve arkadaşlarının yaptığı çalışmanın bulgularına dayanarak temiz hava sahasının kalp hastalıkları sebebiyle yapılan hastane girişi üzerindeki etkisini gösteriyor. Bowling Green'deki temiz hava uygulamasından önce hastaneye başvuru oranı çok daha yüksektir; oranlar yasaktan sonraki her bir yıl düşmüştür. Bu ani değişikliğin sigara yasağı zamanında olmadığına dikkat ediniz. Zamanla gelen aşamalı değişiklik bir müdahalenin etkisini yorumlamayı daha da zorlaştırıyor. Bununla birlikte araştırmacılar kalp hastalıklarının oluşmasının da iyileşmesinin de zaman aldığına dikkat çektiler. Dolayısıyla daha az sigara dumanına maruz kalmanın etkisinin daha belirgin hâle gelmesi, verilerde de gözlendiği gibi, aylar alır.

Koroner kalp hastalığına bağlı hastane başvurusunda sigara yasağından sonraki ilk yılın sonunda %37, üç yıl sonra ise %47'lik bir azalma olmuştur. Genel olarak, müdahalenin etkisinin gecikmeli olarak görülmesi olasılığı olan durumlarda, müdahalenin etkisi hakkında alternatif açıklamaların sayısı da artar (Shadish ve ark., 2002). Bu olaya dair olası bir alternatif açıklama ise Bowling Green'deki kalp hastalıklarına bağlı hastane girişi sayısında azalma olmuş

olması olasılığıdır; bu azalmanın nedeni, daha sigara yasağı yokken bile toplumun dikkatinin yoğun şekilde diyet, egzersiz, ve sigaranın kalp hastalıkları üzerindeki etkilerine yönelmiş olması olabilir.



Bazen araştırmacılar kesintili zaman dizileri ve diğer yarı deneysel desenlerin sonuçlarını, zaman dizisi verilerini gözle denetleyerek yorumlayabilirler. Diğer durumlarda ise sonuçları yorumlamak için daha gelişmiş istatistiksel analizlere ihtiyaç duyulabilir (örneğin; Michielutte, Shelton, Paskett, Tatum & Velez, 2000). Daha fazla bilgi için Shadish ve arkadaşlarının (2002) kitabına bakınız; *Experimental and Quasi-Experimental Designs for Generalized Causal Inference.*

Basit kesintili zaman dizileri deseninde iç geçerliğe yönelik tehditleri incelerken araştırmacılar *tarih etkisine* özellikle dikkat ederler (Shadish ve ark., 2002). İnsanların öz bildirime dayanan karakter gücü durumunda, kişilerin güçlerini yükseltebilecek 11 Eylül saldırısı dışında herhangi bir olay ya da deneyim var mıydı? Peterson ve Seligman (2003), ankette ölçülen merak ve dürüstlük gibi diğer güç'lerin zaman içinde değişmemiş olmasının nedeninin, dinsel güçlerin terörist saldırılardan özellikle etkilendiğini düşündürdüğünü belirtmiştir.

Mevsimsel farklılıklar da dâhil olmak üzere, döngüsel bir örüntü izleyen etkiler, kesintili zaman dizileri deseninde nedensel yorumlara yönelik tehdittir (Cook & Campbell, 1979). Yılın zamanı, Peterson ve Seligman'ın (2003) ölçtüğü dinsel güç için yapılan kesintili zaman analizlerini etkilemiş olabilir. 11 Eylül saldırılarını izleyen haftada ve aylarda, Kefaret Günü, Ramazan, Şükran Günü, Hanuka ve Noel Günü gibi dinsel olan ve olmayan bayramlar kutlandı. Dinsel gücün artmasının nedeni, terörist saldırıların etkisinden ziyade, katılımcıların o mevsime özgü bu bayramların geldiğini gözlemiş olmaları da olabilir. Bu hipotezi test etmenin bir yolu gelecek sene tekrar veri toplayıp bu aylarda (yani Eylül-Aralık aylarında) dinsel gücün artıp artmadığını ölçmek olur. Eğer öyleyse bu mevsimsel değişiklik, Peterson ve Seligman'ın bulgularına alternatif bir açıklama olur.

Basit kesintili zaman dizileri deseninde iç geçerliğe yönelik tehditlerden birisi de araç düzeni olabilir (Shadish ve ark., 2002). Yeni programlar ya da yeni sosyal politikalar başlatıldığı zaman, örneğin bilgi toplamak için kullanılan işlemlerde ya da kayıtların tutulma yöntemlerinde de değişiklikler olur. Cinayeti azaltmaya yönelik bir program, yetkililerin belirli cinayetleri tekrar tanımlamasına ya da bir cinayeti gözlemlerken ve suç faaliyetlerini bildirirken daha dikkatli olmasına neden olabilir. Bununla birlikte, araç düzeni ile ilgili bir tehdit olasılığının olması için, araç düzenindeki değişikliklerin yapılan müdahale ile *tam olarak aynı zamanda* olduğunu göstermek gerekir (Campbell & Stanley, 1966).

Aynı kişiler üstünde hem müdahale öncesi hem de sonrası bir çok gözlem yapıldığı zaman, araştırmacılar iç geçerliğe yönelik tehditlerin olgunlaşma, sınama etkisi ve gerilemeye bağlı olabileceğini de göz önüne almalıdır. Çoğu durumda, çok yönlü gözlemlerin varlığı, araştırmacılara bu tehditleri zaman

dizilerinde olası alternatif açıklamalar olarak eleme olanağı sağlar. Örneğin, olgunlaşmaya bağlı etkiler müdahale ile tam olarak aynı anda belirgin bir kesinti göstermek yerine yavaş yavaş ilerleyebilir. Benzer bir şekilde araştırmacılar müdahale öncesinde zaman dizilerini sınama etkisi ya da gerileme tehditlerine dair herhangi bir bulgu olup olmadığını inceleyebilir.

Basit kesintili zaman dizileri deseninde dış geçerliğe yönelik tehditler dikkatli bir şekilde incelenmelidir. Müdahaleden önce yapılan davranış gözlemleri bir çok testle yapıldığı zaman, müdahale etkisinin bu bir dizi testi alan kişilerle sınırlı kalması olasılığı yüksektir. Dahası, kesintili zaman dizileri deseninde genellikle seçkisiz olarak seçilmemiş sadece bir grup test edilir. Desenin bu yönü, sonucun ancak çalışmada yer alan kişilere benzer özelliklere sahip olanlarla sınırlı kalması olasılığını ortaya çıkarır.

Denk Olmayan Kontrol Gruplu Zaman Dizileri Deseni

Denk olmayan kontrol gruplu zaman dizileri deseninde, araştırmacı hem müdahale hem de karşılaştırma grubunda, müdahaleden önce ve sonra bir dizi gözlem yaparlar.

Temel Kavram

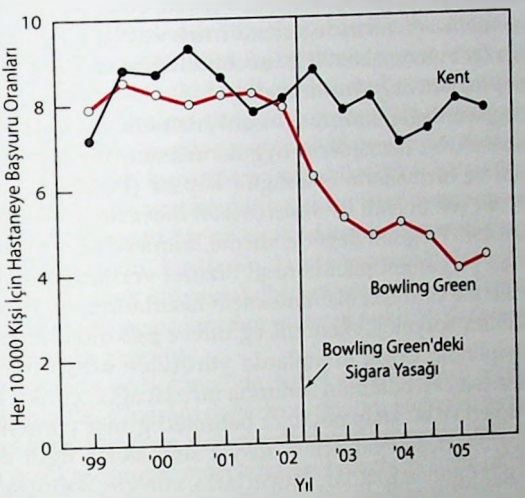
Kesintili zaman dizilerinin iç geçerliği denk olmayan kontrol grup deseninde açıkladığımız işlemlere ilaveten bir kontrol grubu eklenerek fazlasıyla güçlendirilebilir. **Denk olmayan kontrol gruplu zaman dizileri deseninde** araştırmacı, müdahale grubuna denk olan, müdahalenin deneysel gruba verilmesinden önce ve sonra benzer şekilde gözlemlerin yapıldığı bir grup bulmak zorundadır. Desen şu şekilde sembolleştirilebilir:

$$\begin{array}{c} O_1\ O_2\ O_3\ O_4\ O_5\ X\ O_6\ O_7\ O_8\ O_9\ O_{10} \\ \hline O_1\ O_2\ O_3\ O_4\ O_5 \qquad O_6\ O_7\ O_8\ O_9\ O_{10} \end{array}$$

Önce de belirtildiği gibi kesik çizgiler kontrol ve deney gruplarının seçkisiz olarak atanmadığını gösteriyor. Denk olmayan kontrol gruplu kesintili zaman dizileri deseni araştırmacılara tarih etkisi tehdidini ortadan kaldırma olanağı tanır. Örneğin, Khuder ve arkadaşları sigara yasağının Bowling Green-Ohio'daki etkisini incelerken, bir yandan da Bowling Green'den 150 mil uzakta olan Kent-Ohio'daki hastaneye yatış kayıtlarını incelediler. Denk olmayan kontrol grubu olarak Kent'i seçme nedenleri, evren, yaş ve cinsiyet dağılımdaki benzerliğin dışında, daha da önemlisi Kent bölgesindeki yetkililerin sigara yasağını başlatmamasıydı. Şekil 10.7 bu iki şehirdeki hastaneye yatanları gösteriyor. Kalp hastalığı şikâyetiyle yapılan hastane başvuruları Bowling Green'de sigara yasağı sonrası azalırken, bu etki Kent-Ohio'da gözlemlenmemiştir.

Araştırmacılar denk olmayan kontrol gruplu zaman dizileri deseni kullanarak iç geçerliğe yönelik birçok tehdidi ortadan kaldırabilirler. Örneğin, Bowling Green'de gözlenen hastaneye yatışların toplumun diyet, egzersiz ve sigaraya yönelttiği dikkat nedeniyle azalmış olabileceği olasılığını dışlayabiliyoruz çünkü bu dikkat Kent'te yaşayan insanları da etkilerdi. Makul olmak gerekirse, Bowling Green'deki hastaneye yatış oranlarının düşmesine alternatif bir

ŞEKİL 10.7 1999-2005 yılları arasında her 10.000 kişi için kalp şikâyeti nedeniyle yapılan hastane başvurularının Bowling Green ve Kent arasındaki 6 aylık (Ocak-Haziran, Temmuz-Aralık) ortalamalarının müdahale ve denk gruplar arasında karşılaştırılması. Veriler, Khuder ve arkadaşların'dan (2007) alınmıştır.



açıklama, Kent yerine Bowling Green'deki insanları etkilemesi gerekir ve sigara yasağıyla aynı zaman diliminde ortaya çıkması gerekir (yani, seçimin ilave etkisi). Gördüğünüz gibi, müdahaleden önce ve sonra birçok gözlem yapılmışsa ve iyi seçilen bir kontrol grubu varsa, müdahale etkisinin dışında alternatif başka açıklamalar yapmak zor olur; işte bu özellik, denk olmayan kontrol grup desenini güçlü bir yarı deneysel desen yapar.

PROGRAM DEĞERLENDİRME

- Program değerlendirme, insanî hizmet kurumlarının etkililiğini ölçmek ve yöneticilere, verdikleri hizmetler hakkında geri bildirimde bulunmak için kullanılır.
- Program değerlendiriciler, sosyal hizmetlerin ihtiyaçlarını, süreçlerini, sonuçlarını ve etkililiklerini ölçer.
- Temel ve uygulamalı araştırmalar arasındaki ilişki karşılıklıdır.
- Toplumun deneye karşı isteksiz olmasına karşın, gerçek deneyler ve yarı deneysel desenler sosyal reformları değerlendirmek için mükemmel yaklaşımlar sağlayabilir.

Üretim yapan bir kuruluş başarıyı yakalar. Eğer bir firma dokunmalı tablet üretiyorsa, başarısını en sonunda tablet satışlarından elde ettiği kârı belirler. En azından teorik olarak kurumun verimliliği ve etkililiği, muhasebe defterini inceleyerek kolayca ölçülebilir. Bununla birlikte, giderek artan bir şekilde, farklı tip kurumlar toplumda önemli bir rol oynar. Bu organizasyonlar üründen

ziyade hizmet sağladıkları için, Posavac (2011) onları insanî hizmet kurumları (İHK) olarak adlandırdı. Örneğin, hastaneler, okullar, polis merkezleri ve devlet kurumları acil bakım odalarından yangın önleme denetimine kadar çok çeşitli hizmetler sunar. Bu kurumların kâr etme amaçları olmadığı için, etkililik ve verimlilik açısından değerlendirmek ve birbirlerinden ayırt etmek için farklı yöntemler bulunmalıdır. İHK'nın etkililiğini ölçmenin en kullanışlı yollarından birisi, program değerlendirmedir.

Temel Kavram

Program değerlendirme, insanî hizmetleri değerlendirmek için kullanılan yöntemleri, bu hizmetlerin uygulanmasını, hizmet verilen insanlara olan etkilerini ve hizmetlerin etkililiğini kapsar (Posavac, 2011). Program değerlendirmenin çok önemli *bir amacı, insanî hizmetler etkinliklerine ilişkin geri bildirim sağlamaktır*. Program değerlendirme, kimlere ne tip hizmetlerin sunulacağına, en etkili ve verimli şekilde nasıl hizmet verileceğine dair karar verirken İHK yetkililerine yardımcı olabilmek için tasarlanmıştır. Program değerlendirme siyasal bilim, sosyoloji, ekonomi, eğitim ve psikolojiden faydalanan bütünleştirici bir disiplindir. Doğal ortamlarda yürütülen araştırmalardaki program değerlendirmesini bu bölümün sonunda tartışacağız, çünkü belki de en geniş ölçekli prensipleri ve bu kitap boyunca betimlediğimiz yöntemleri yansıtmaktadır.

Program değerlendiriciler, sosyal servislerle ilgili dört tip araştırma sorusunu göz önünde tutarlar: İhtiyaçlarla, süreçle, sonuçla ve etkililikle ilgili sorular (Posavac, 2011). *İhtiyaçların* değerlendirilmesi, bir kurumun hizmet verdiği insanların karşılanmamış ihtiyaçlarını belirler. Örneğin, kent yönetimine toplumdaki yaşlı insanların eğlence faaliyetleri ile ilgili program önerildiğini düşünün. Öncelikle kent yönetimi yaşlı kişilerin bu tip bir programa ihtiyaçları olup olmadığı ya da isteyip istemediklerini belirlemek ister. Eğer yaşlılar böyle bir program istiyorlarsa, yönetim hangi tip programların onlar için en ilgi çekici olacağını bilmek ister. Anketler, ihtiyaçları belirleme çalışmalarında geniş ölçüde kullanılır. Yönetim, anketlerden elde edilen bilgiyi ne tip programlar önereceğine yardımcı olması için kullanır.

Bir kez program başlatıldıktan sonra, program değerlendiriciler oluşturulan *süreç* hakkında sorular sorabilir. Gözlemsel yöntemler program sürecini ölçmek için genelde kullanışlıdır. Programlar her zaman planlandığı şekilde uygulanamaz ve bir program uygulandığı zaman ne yapıldığını bilmek gerekir. Eğer yaşlılar düzenlenen eğlence programı içindeki aktiviteleri yapmıyorlarsa, programın yetersiz kaldığı düşünülebilir. Süreçle ilgili sorulara cevap veren, sürecin nasıl işlediği hakkında bilgi veren bir değerlendirme, mevcut programın güçlendirilmesi için yetkililere düzenleme yapma olanağı sağlar (Posavac, 2011).

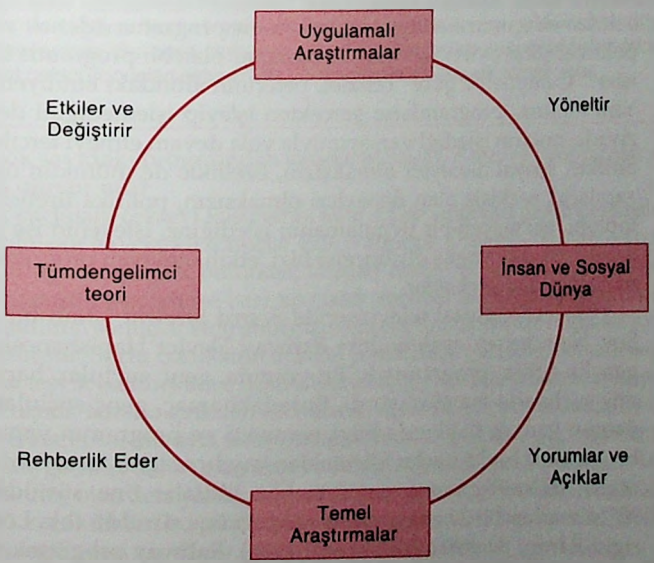
Bir program değerlendiricinin soracağı bir sonraki soru grubu, *sonucu* kapsar. Program belirlenen hedefleri karşıladı mı? Örneğin, şimdi yaşlılar daha fazla eğlenme aktiviteleri yapıyorlar mı ve bu aktivitelerden memnunlar mı? Bu aktiviteleri, diğer aktivitelere tercih ediyorlar mı? Mahallede suç işlenmesini engellemek ve güvenliği arttırmak amacıyla hazırlanan "mahalleyi gözleme programı"nın (neighborhood-watch program) uygulanmasından sonra hırsızlık ve saldırıların azalıp azalmadığı değerlendirilebilir. Sonuçları değerlendirmek için, Bölüm 4'te açıklananlara benzer arşiv verileri kullanılabilir. Örneğin,

çeşitli suçların sıklığını belgelemek için polis kayıtlarını inceleme, bu programın ne ölçüde etkili olduğunu değerlendirmenin bir şeklidir. Sonuçları değerlendirmek için, doğal ortamda kullanılan hem deneysel hem de yarı-deneysel yöntemler de kullanılabilir. Örneğin, bir okul reform programı, denk olmayan kontrol grubu deseni kullanılarak değerlendirilebilir. Bu amaçla bir değerlendirici, öğrencilerin performansını, birisi programın uygulandığı diğeri ise uygulanmadığı okul olmak üzere, iki farklı okulu karşılaştırarak değerlendirebilir.

Değerlendiricilerin sormaları gereken son soru, programın *etkililiği* hakkındadır. Etkililik hakkındaki sorular, çoğunlukla programın maliyetiyle ilgilidir. Bir hükumetin veya diğer kurumların yapabilecekleri hizmetler arasında çoğunlukla bir seçim yapma zorunluluğu vardır. Eğer programın devamı, onun nasıl geliştirileceği, alternatif bir program denenip denenmeyeceği veya programın verdiği hizmetlerin kısıtlanıp kısıtlanmayacağı konularında verilere dayanan bir karar vermek istiyorsak, bir programın ne kadar başarılı olduğu (sonuç değerlendirmesi) ve programın maliyeti (etkililik değerlendirmesi) hakkında bilgi edinmemiz gerekir.

Bu bölümün başında ve Bölüm 2'de, temel ve uygulamalı araştırmalar arasındaki farkları betimledik. Program değerlendirme, muhtemelen uygulamalı araştırmanın uç bir şeklidir. Program değerlendirmenin amacı pratiktir, teorik değildir. Böyle olmakla birlikte, pratik amaçların en belirgin olduğu bir bağlamda bile, temel ve uygulamalı araştırmalar arasında bir ilişki kurulabilir. Bu ilişkinin bir modeli, Şekil 10.8'de görülmektedir. Buradaki fikir şudur: Her bir araştırma alanı, döngüsel şekilde diğer alanı destekler. Yani, temel araştırma bize, doğadaki belirli düzenlilikleri ifade eden temel bilimsel prensipleri sağlar. Bu prensipler, işledikleri

ŞEKİL 10.8 Temel ve uygulamalı araştırmalar arasındaki karşılıklı ilişkiyi gösteren model (Solomon, 1987 s. 444'ten alınmıştır)



varsayılan karmaşık ve "kirli" dünyada incelendikleri zaman yeni karmaşık ilişkiler görülür ve yeni hipotezler kurulur. Bu yeni karmaşık ilişkiler daha sonra, tekrar gerçek dünyada denenmeden önce, laboratuvarda test edilir ve değerlendirilirler.

Ellen Langer'ın çalışması, bu döngüsel ilişkinin somut bir örneğidir (bk. Salomon, 1978). Huzurevine girdikten sonra yaşlıların sağlık durumlarında bir düşüş buldu (bk. Langer, 1989; Langer & Rodin, 1976). Bu doğal gözlemler, onu bir farkındalık teorisine (mindfulness theory) götürdü ve teorisini kontrollü deney koşullarında test etti; bu teorinin daha genel bilişsel gelişim, eğitim, ve sağlık teorilerini test etmeye yarayacak sonuçları da vardı (bk. örneğin; Hsu, Chung, & Langer, 2010; Langer, 1989, 1997; Langer & Piper, 1987). Teorisi, uygulamalı araştırmalarında Langer'a rehber oldu – yeni model huzur evleri tasarlamada. Huzur evlerinde verilen bakımda yapılacak değişikliklerin, yaşlıların sağlığı ve mutluluğu üzerindeki pratik etkilerinin test edilmesi, onu farkındalık teorisinde değişiklikler yapmaya yöneltti.

Campbell'e (1969) göre, sosyal deneylere katılan kamu görevlileri için çözümün öneminden ziyade problemin önemini vurgulamak daha önemlidir. Belirli bir "toptancı" çözümde (ki çoğu durumda başarı şansı düşüktür) ısrar etmek yerine, görevliler en açık değerlendirmeye izin veren bir tarzda reform yapmaya hazır olmalı ve eğer bir çözüm başarısız olursa, farklı bir çözüm denemeye hazırlanmalıdırlar. Bir başka ifadeyle, kamu görevlileri toplumun sorunlarını teşhis etmek ve etkili çözümler bulmak için deneysel yöntemi kullanmaya hazır olmalıdır (bk. Kutu 10.2).

Campbell'in (1969) sosyal reformlar ile deneysel yöntemlerin düzenli şekilde buluşturulması fikri politika üreticilerini etkiledi ama henüz yeterince kullanıldığı söylenemez (bk. Berk ve ark., 1987). Çeşitli tipte sosyal reformlar düzenli şekilde uygulanmakla birlikte, bu reformların başarısı nadiren değerlendirilmektedir. Campbell 1969 yılında, politik atmosferin bilimsel değerlendirmeyi kabul etmeye hazır olmadığını öne sürdü; muhtemelen hâlâ öyledir. Etkililiğini belirlemek için test edilmesi gereken bir programa ödenek verilmesi oylansa, politikacıların oyu nasıl olur? Ve başarısız olan bir programla ilişkili olmayı kim ister? Campbell'e göre "cehalet, pelerinin altındaki emniyettir" (ss. 409–410). Yani toplum, programların gerçekten işleyip işlemediğini değerlendirmekten ziyade, onların işlediği varsayımıyla yola devam etmeyi tercih ediyor. Bununla birlikte, sosyal deneyler olmaksızın, özellikle de mümkün olan her durumda yapılacak seçkisiz alan deneyleri olmaksızın, politika üreticileri ve genellikle toplum, işlemeyen bir uygulamanın işlediğine, işleyenin ise işlemediğine inanabilir. Bu gibi yanlış düşünceler bizi, etkili olmayan programlara para ve kaynak ayırmaya sevk eder.

1980'lerde, ulusal televizyonda *Scared Straight* isimli bir şov yayımlandı. Şov, New Jersey eyaletindeki Rahway Devlet Hapishanesinde yapılmış bir gençlik eğitim programıydı. Programda, genç suçlular, hapishanedeki seçilmiş suçlularla tanıştırılıyordu. Buradaki amaç, genç suçlulara hapishane hayatının gerçeği hakkında bilgi vermekti ve programın yapımcısı bu suretle, bu gençleri başka suçlar işlemekten caydıracağını umuyordu. Bu programın etkisi hakkında kanıta dayanmayan iddialar öne sürüldü; örneğin, %80, 90'lara varan bir başarı oranına ulaştığı öne sürüldü (bk. Locke, Johnson, Kirigin-Ramp, Atwater & Gerrard, 1986). Rahway programı, ülke genelindeki

KUTU 10.2

TOPLUM DENEYLERİNİN SOSYAL VE POLİTİK GERÇEKLERİ

Sağlık hizmeti vermek zorlu, çok yatırım ve çaba isteyen bir hizmettir (belki de bu ifadeyi politik kampanyalarda duydunuz). Sağlık hizmeti verenler, yaşlı evrenindeki kalp, diyabet gibi kronik hastalıkların tedavisiyle ilişkili tedaviler esnasında birçok zorlukla karşılaşırlar. Ayrıca, günümüzde insanlar kanser, AIDS, kalp yetmezliği gibi öldürücü hastalıklarla yıllarca yaşayabiliyorlar. Mevcut tedaviler bu kronik hastalıkları tedavi etmiyor, hastaların sıkıntılarını olabildiğince hafifletmeye çalışıyor. Ama sağlık bakımı programları etkili oluyor mu? Programlar, planlandıkları gibi uygulanıyor mu? Hangi programlar daha etkili ve uygun maliyetlidir? Program değerlendirme alanındaki sorular bunlardır.

ABD'nin 65 yaş üzerindekiler için millî sağlık sigortası programı Medicare, Doylestown, Pennsylvania'da, Health Quality Partners (HQP; Klein, 2013) adı verilen bir sağlık bakım programını değerlendirdi. Medicare 1967 yılından bu yana yenilikçi sağlık bakımı geliştirmek üzere hazırlanmış olan çok sayıda araştırma projesini destekledi. Gerçek bir deney, HQP sağlık bakım programını test etmek için, kronik hasta olan yaşlıları seçkisiz olarak tedavi ve kontrol gruplarına atadı. Tedavi koşulunda, eğitimli bir hemşire düzenli olarak ev ziyaretleri yaptı. Bu ziyaretler esnasında hemşire, hastaların sağlığını ve ilaçlarını kontrol etti ve onların sorularını cevapladı. Bu basit fikrin ve uygulamanın dikkat çekici sonuçları oldu: HQP, hastaneye yatma oranını %35 düşürdü ve Medicare'ın maliyetlerini %22 azalttı.

Bu başarısına rağmen Medicare, HQP programını kaldırmaya karar verdi. Deney kötü olduğu için veya sonuçlar tartışmalı olduğu için değil, ABD'de sağlık hizmeti sağlamayla ilişkili gerçekler yüzünden kaldırıldı. Medicare yetkilileri, HQP'nin böyle bir hizmeti daha geniş kitlelere sağlamasının mümkün olmadığı sonucuna vardı. Bu, bir hizmetin "genişletilebilir" olup olmadığı meselesidir – yani, küçük bir gruba verilen bir hizmetin, geniş kitlelere yaygınlaştırılabilme derecesi. Aynı zamanda, HQP hastanelerde hâlen verilmekte olan bakım modeliyle uyumlu olmadığı için, programın genişletilmesinin ekonomik ve politik gerçekleri daha da büyümektedir (Klein, 2013). Medicare yetkilileri, bakımın kalitesinden ziyade teşvik primlerini incelemek için hazırlanmış araştırma projelerine daha fazla odaklanmaya karar verdiler (Kongre Bütçe Bürosu, Ocak, 2012). Bu nedenle, program kaldırıldı.

Doğru karar bu mudur? ABD'deki büyük, karmaşık sağlık hizmeti verme problemi çerçevesinde, tek bir "doğru" karar yoktur ve insanlar bu konuda, çoğunlukla keskin bir görüş ayrılığı içindedirler. *Bununla birlikte, sosyal bilimciler, karar verme sürecini bilgilendirecek araştırmaları istemeye ve iyi kontrol edilmiş araştırma projelerinden elde edilen bilgilerin toplumsal sorunlara uygulandığına inandırmak için çalışmaya devam etmelidirler.*

benzer çok sayıdaki programdan sadece birisidir. Ama bu programlar gerçekten etkili oluyor mu?

Çocuk suçlulara cezaevi koşullarını tanıtma programları çeşitli çalışmalarla değerlendirilmiştir ama farklı sonuçlar elde edilmiştir. Örneğin, bazıları bu programların suçluluğu azalttığını bulurken, bazıları deney ve kontrol grupları arasında fark bulmamıştır; bazıları ise bazı tip çocuk suçlular arasında suç oranını *arttırdığını* düşündüren sonuçlar elde etmiştir. Neticede, bu uyumsuz bulgular şu şekilde yorumlanmıştır: Suçlu çocukların arasına yeni katılan çocuklar, daha deneyimliler tarafından teşvik edilmekte ve desteklenmektedir; işte bu etki cezaevi koşullarıyla deneysel olarak temasa getirilen çocuklarda daha da güçlenmektedir. Dolayısıyla bu grup, cezaevi koşullarıyla tanıştırıldığı zaman, suç eğilimi güçlenmektedir. Öte yandan, suç geçmişi daha uzun olan ve ait olduğu suçlu grubu arasında belirli bir statü kazanmış olan deneyimli çocuk suçlular, cezaevini bir tehdit olarak algılıyorlar, çünkü kazandıkları bu statüyü kaybetme endişesine kapılıyorlar. Dolayısıyla, suç işleme eğilimleri zayıflıyor (bk. Locke ve ark., 1986).

Bu önemli sosyal programın etkililiğini değerlendirme girişimleri, değerlendirme araştırmalarının doğasında olan zorlukları iyi bir şekilde ortaya çıkardı: Katılımcıları gruplara seçkisiz atama zorluğu, yöneticileri deneysel işlemlere uygun davranmaya ve işbirliği yapmaya ikna etme ve değerlendirme süreci içinde bazı katılımcıların kaybedilmesi sorunuyla başa çıkma. Böyle olmakla birlikte, sağlam deneysel yöntemlere dayanan program değerlendirme, politika üreticilerinin her düzeyde (kurum, toplum, kent, eyalet, federal yapı), sosyal sorunlara farklı müdahale seçenekleri arasında karar vermelerine yardımcı olacak bilgi sağlar. Kaynaklar daima kısıtlı olduğu için, en verimli şekilde kullanılmaları gerekir. Öyle umuyoruz ki sizin araştırma yöntemleri bilginiz, deney yapmanın toplum içindeki rolü konusunda devam etmekte olan tartışmalara, bilgili ve yapıcı bir şekilde katılmanızı sağlayacaktır.

ÖZET

Doğal ortamlarda deney yapmak, psikoloji laboratuvarlarında deney yapmaktan birçok yönüyle farklıdır. Doğal ortamlarda deney yapmanın nedenleri, laboratuvar bulgularının dış geçerliğini test etmek ve insanların yaşadıkları ve çalıştıkları koşulları iyileştirmeyi amaçlayan "müdahale"lerin etkilerini değerlendirmektir.

Çoğu sosyal bilimciye göre toplum, sosyal reformlara deneysel yaklaşımı desteklemelidir –bu yaklaşım, yeni programların etkililiğini en açık şekilde değerlendirme olanağı sağlar. Çoğu durumda (örneğin, mevcut kaynaklar yetersiz olduğu zaman), müdahale ve kontrol gruplarına katılımcıları seçkisiz atayan gerçek deneyler önerilir. Böyle olmakla birlikte, eğer gerçek bir deney uygulanamıyorsa, ikinci en iyi seçenek yarı deneysel yöntemlerdir. Yarı deneysel yöntemler gerçek deneylerden farklıdır çünkü bir deneysel sonuç için az sayıda alternatif hipotez kontrol edilir. İç geçerliğe yönelik spesifik bir tehdit kontrol edilmediği zaman, deneyci, durumu mantıken inceleyerek ve ek kanıtlar toplayarak, iç geçerliğe yönelik bu tehdidi elemeye çalışmalıdır.

Özellikle güçlü bir yarı deneysel desen, denk olmayan kontrol gruplu desendir. Bu desen, arttırıcı etkilerle ilişkili olanlar hariç, iç geçerliğe yönelik önemli tüm tehditleri kontrol eder: (1) seçim ve tarih etkisi, (2) seçim ve olgunlaşma etkisi, (3) seçim ve araçlandırma etkisi, (4) diferansiyel istatistiksel gerilemeye bağlı tehditler. İç geçerliğe yönelik başlıca tehditlere ilaveten, bir araştırmacı katılımcı grupları arasındaki haberleşmenin neden olabileceği kirlenmeye karşı da dikkatli olmalıdır. Deneyci beklentisi etkileri sorunu (gözlemci yanlılığı), dış geçerlik sorunları ve Howthorne etkisi de dâhil yenilik etkileri, laboratuvarda veya alanda olması fark etmeksizin, bütün deneylerin potansiyel sorunlarıdır.

Bir bağımlı ölçümdeki değişmeleri bir müdahaleden önce ve sonra gözlemek mümkün olduğu zaman, bir basit kesintili zaman dizileri deseni uygulanabilir. Bu deseni kullanan araştırmacı, müdahalenin yapılmasıyla birlikte zaman dizilerinde ortaya çıkan ani bir değişim (kesinti) arar. Bu desende iç geçerliğe

yönelik en önemli tehdit, tarihtir – zaman dizilerindeki değişmeye, müdahalenin değil de başka bazı olayların neden olmuş olması demektir. Araçlandırma da bir sorun olabilir, özellikle bir tedavi (müdahale), kayıtların tutulma veya verilerin toplanma şeklinde değişmelere yol açan bir sosyal reform tipini temsil ettiği zaman bu sorun görülür. Deneysel gruba olabildiğince benzer olan bir kontrol grubu dâhil ederek, bir basit zaman dizileri deseninin iç geçerliği güçlendirilebilir. Denk olmayan bir zaman dizileri kontrol grubu, örneğin, birçok muhtemel tarih tehdidini kontrol eder.

Doğal ortamlarda araştırma yapmanın özellikle önemli bir amacı, program değerlendirmedir. Psikologların dışındaki (örneğin; eğitimciler, siyaset bilimciler, ve sosyologlar gibi) profesyoneller çoğunlukla bu sürece katılırlar. Program değerlendirme tiplerinden bazıları ihtiyaçların, süreçlerin, sonuçların ve etkililiğin, değerlendirilmesidir. Program değerlendirmenin üzerindeki belki de en ciddi kısıtlılık, programın içinde gerçekleştiği politik ve sosyal gerçeklerdir. Çoğunlukla üstesinden gelinmesi gereken bir engel, kamu görevlilerinin sosyal reformların değerlendirilmesine karşı isteksizlik göstermeleridir. Böyle olmakla birlikte sosyal bilimciler, program değerlendiricilerini insan hizmetleri kurumlarında görev almaya çağırmaktadırlar. Bu çağrıya cevap vererek biz, en çok ihtiyaç duyanlara en etkili hizmeti verecek şekilde toplumun değişmesine yardım edebiliriz.

TEMEL KAVRAMLAR

iç geçerliğe tehditler 310
olgunlaşma 310
sınama etkisi 310
tarih 310
araç düzeni 311
regresyon 311
katılımcı kaybı 312
seçim 312
kirlenme 314
yenilik etkileri 315
yarı-deneysel desenler 317
denk olmayan kontrol gruplu desen 318
basit kesintili zaman dizileri deseni 326
denk olmayan kontrol gruplu zaman dizileri deseni 330
program değerlendirme 332

GÖZDEN GEÇİRME SORULARI

1 Doğal ortamda deney yapmanın özellikle önemli olan iki nedenini açıklayın.
2 Laboratuvar deneyleri ile doğal ortamda yapılan deneylerin kontrol, dış geçerlik, amaç ve sonuçlar açısından nasıl farklı olduklarını açıklayın.
3 Gerçek deneylerin ayırt edici üç özelliğini açıklayın ve bu özelliklere dayanan bağımsız değişken hakkında neler söylenebileceğini belirleyin.
4 Doğal ortamlarda deney yaptıkları zaman araştırmacıların üstesinden gelmek zorunda oldukları engelleri açıklayın.
5 Tüm katılımcılara deneysel müdahale yapılırken, araştırmacının katılımcıları deneysel koşullara seçkisiz atamasına izin veren iki işlemi açıklayın.
6 Kirlenme meydana geldiği zaman bir kontrol grubundaki katılımcıların tepki verebilecekleri üç durumun sonuçlarını betimleyin.

İSTANBUL TİCARET ÜNİVERSİTESİ
KÜTÜPHANESİ

7 Howthorne etkisi dâhil yenilik etkilerinin, bir araştırmacının bir deneysel müdahalenin sonuçlarına dair yorumlarını nasıl etkileyebileceğini açıklayın.
8 Dış geçerlik en iyi nasıl test edilebilir?
9 Denk olmayan kontrol gruplu desende ön test yapmanın neden zorunlu olduğunu açıklayın.
10 Denk olmayan kontrol gruplu desende, iç geçerliğe yönelik bir tehdidin nasıl kontrol edileceğini açıklayın ve bu desende kontrol edilmeyen bir tehdit söyleyin.
11 Denk olmayan kontrol grup deseninde uygulama ve kontrol gruplarının, her iki grubun ön test puanları aynı olsa bile, denk oldukları sonucuna neden varamayacağımızın iki nedenini açıklayın.
12 Denk olmayan kontrol gruplu desende, iç geçerliğe yönelik tarih tehdidi ile "yerel tarih etkisi" denilen etki arasındaki farkı açıklayın.
13 Bir basit kesintili zaman dizileri deseninde, bir müdahalenin etkili olduğunun başlıca kanıtları nelerdir?
14 Bir basit kesintili zaman dizileri desenine bir denk olmayan kontrol grubu eklemenin, desenin iç geçerliğine yönelik tehdidi nasıl azaltacağını açıklayın.
15 Program değerlendirmede değerlendiricilerin, dört sorunun her birisiyle edinmek istedikleri bilginin türünü açıklayın.

DÜŞÜNDÜRÜCÜ SORULAR

1 Multimedya ile öğretimin etkili olup olmadığını belirlemek için bir yarı deneysel yöntem kullanıldı. Psikolojiye giriş dersinin iki bölümünü, her ikisi de öğleden sonra olmak üzere, aynı öğretmen öğretti. Bir bölümde (uygulama grubu) öğretmen multimedya öğretimi kullandı. Diğer bölümde ise aynı materyali kullanmakla birlikte, multimedya kullanmadı. Öğrencilerin, hangi gruba kaydedildiklerinden haberleri yoktu ama bu gruplara seçkisiz atanmadılar. Öğrencilerin bilgisi, kapsamlı bir psikolojiye giriş testinin iki formuyla değerlendirildi. Bu kapsamlı test, iki grupta verilen dersin etkililiğini karşılaştırmak için kullanılabilecek güvenilir ve geçerli bir test olarak kabul edilebilir. Her iki bölümdeki öğrencilere dersin ikinci günü ön test, son günü ise son test uygulandı; bu testlerde, testin farklı formları uygulandı.
   A Bu araştırmada hangi yarı deneysel desen kullanıldı?
   B Öğretmen başlangıçta, yarı deneysel değil gerçek bir deney yapmayı düşündü. Eğer siz multimedya öğretiminin etkililiğini test etmek için bir gerçek deney yapılmasından yana görüş belirtseydiniz, seçkisiz atamanın doğruluğu hakkında nasıl bir eleştirel yorum yapardınız.
   C Öğretmenin kullandığı yarı deneysel desenin, neden sadece multimedya ile ders alan öğrencileri test etmesinden daha etkili olduğunu açıklayın. Bu araştırmada kontrol edilmiş olan, ama eğer sadece multimedya ile ders alan öğrenciler test edilmiş olsaydı kontrol edilemeyecek olan bir iç geçerlik tehdidi bulun.
2 Bir psikolog, boşanmanın erkekler, kadınlar ve çocuklar üzerindeki etkilerini ele alan bir kitap yayımladı. Boşanmanın, boşanmadan 10 yıl sonraki etkileriyle ilgilendi. Kadınların yarısının, erkeklerinse üçte birisinin, boşanmadan 10 yıl sonra bile hâlâ birbirlerine büyük öfke duyduklarını belirledi. Erkeklerin ve kadınların yarısı mutlu olduklarını ifade etmiş olmakla birlikte, kadınların %25'i ve erkeklerin %20'si hayatlarını "yoluna koyamadıklarını" bildirdiler. Boşanmış ailelerin sadece %10'unda hem eski kocalar hem de kadınlar 10 yıl sonra mutlu, ve hayatlarından memnun yaşıyorlardı. Nihayet, boşanmış ailelerin çocuklarının yarısından fazlası, yetişkinliğe başarısız ve kendini değersiz gören kadınlar ve erkekler olarak adım attılar. Bu bulgular California Marin County'de (Çoğunlukla eğitim düzeyi yüksek ve varlıklı insanların yaşadığı bir banliyöde) yaşayan 60 boşanmış çift ve onların

131 çocuğu üzerinde yapılmış olan ve 15 yıl sürmüş bir araştırmadan alınmıştır).

A Bu araştırmadaki bulguların (örneğin; mutluluk, doyum) boşanmanın etkilerine bağlı olduğu sonucuna varmak ikna edici midir? Neden evet, neden hayır?

B Bu araştırmanın, denk olmayan kontrol gruplu zaman dizileri deseniyle nasıl yapılabileceğini açıklayın. Araştırmacının, boşanmanın etkilerine dair nedensel bir ilişki kurabilmesi için nasıl bir veri örüntüsü gerekir?

C Bu araştırma hakkında iki eleştirmen yorumu var. Birisi, dış geçerliği olmadığını söyleyerek eleştiriyor. İkinci eleştirmen ise dış geçerliğin olup olmadığını bilmediğini söylüyor. Siz hangi eleştirmene katılıyorsunuz?

3 Büyük bir şehrin polis teşkilatı, polis memurlarının yasa değişikliklerinden haberdar olmasını sağlamak için iki ayrı yaklaşım arasında tercih yapmak zorundaydı. Teşkilatın zeki bir yöneticisi, bu iki yaklaşımı bir araştırmayla test etmeye karar verdi. Gerçek bir deney yapılmasını istedi ve altı aylık bir süre için iki programın her birisine seçkisiz olarak 30 memur atadı. Bu sürenin sonunda, iki yaklaşım altında eğitimi başarıyla tamamlayan memurlara, yasa bilgilerini test eden bir final testi verildi. Bu testte, A programını tamamlayan 20 memur, B programını tamamlayan 28 memurdan daha yüksek puan aldı. Yönetici, akıllıca bir tutumla, bu sonuçları iki programın etkililiğinin kesin bir kanıtı olarak kabul etmemeye karar verdi.

A Sadece bu özetteki bilgiyi kullanarak, yöneticinin neden bu kararı verdiğini açıklayın. Yani, yönetici iç geçerliğe hangi tehdidin olduğunu belirledi?

B Her bir gruptan (orijinal gruba atanan 30 kişiden) sadece 20 memur tamamlamış olsaydı, ve Program A lehine hâlâ önemli bir fark kalsaydı, yöneticinin kararı nasıl değişebilirdi?

C Yönetici programlardan birisinin daha etkili olduğunu söyleyebilir mi?

4 Yeni bir fitnes merkezi kuran küçük bir üniversite, fakülte öğrencileri ile çalışanlar için bir sağlık geliştirme programı uygulamaya karar verdi. Programa göre haftada üç defa birer saatlik uygulama yapılacaktı ve bir yarıyıl sürecekti. Bu programın değerlendirilmesine ilişkin aşağıdaki soruları cevaplayın.

A Programın planlanmasında ihtiyaçların değerlendirilmesi nasıl bir rol oynayabilir?

B Program bir defa uygulanmaya başladıktan sonra, programın sonuçlarının değerlendirilmesinin uygun şekilde yorumlanabileceğini güvence altına almak için, program süreci hakkında hangi soruların sorulması yararlı olur?

C Eğer gerçek bir deney mümkün olmazsa, önerilen programın etkililiğini nasıl test edeceğinizi açıklayın.

İleri Alıştırmaların Cevapları

1 Katılımcılardaki iyileşmeyi müdahale değil de başka bir olay açıklayabildiği zaman, tarih bir tehdittir. Örneğin, katılımcılar kendini geliştirme kitapları okumuş, bitkisel ilaçlar denemiş, arkadaşlarıyla veya papazlarla konuşmuş, veyahut potansiyel olarak faydalı "müdahaleler" almış olabilirler. Psikoloğun müdahalesinden ziyade, bu diğer olayların herhangi birisi depresyonun iyileşmesine neden olmuş olabilir.

2 Katılımcılar zamanla doğal olarak değiştiklerinde, olgunlaşma ortaya çıkar. Depresyon hakkında bildiğimiz şeylerden birisi, zaman içinde iyileşme eğiliminde olduğudur. Bu nedenle, katılımcıların iyileşmesi, tedaviden ziyade, zamanla depresyondaki doğal azalma olabilir.

3 Bir testin ilk uygulaması bir sonraki testi etkilediği zaman, test etkisi tehdidi ortaya çıkar. Bu araştırmada katılımcılar, bir depresyon ölçeğinde önceden verdikleri tepkileri hatırlamış olabilirler ve, belki de iyileştiklerini göstermek için, son testte daha hafif depresyon belirten seçeneği işaretlemiş olabilirler (kendilerini daha az depresif hissetmedikleri hâlde). Bir ölçek düşüncelerin, duyguların, ve davranışların zaman içindeki değişmesini ölçmek için kullanıldığı zaman araçlandırma tehdidi ortaya çıkar. Hem ön testte hem de son testte aynı anket kullanıldığı için, bu tehdidin ortaya çıkması olasılığı daha düşüktür.

4 Bir ön test sonucuna göre uç puanlar alan katılımcılar seçildiği zaman, istatistiksel gerileme gözlenmesi mümkündür. Bu araştırmada katılımcılar, depresif oldukları için seçildiler – bir depresyon ölçeğinden yüksek puanlar aldılar. Son testteki daha düşük puanlar, tedavinin etkisini değil, ortalamaya istatistiksel gerilemeyi gösteriyor olabilir.

Düşündürücü Soruların Cevapları

A Bu araştırmada denk olmayan kontrol grup deseni kullanıldı.

B Öğrenciler iki gruba seçkisiz atanmayı haksızlık olarak görebilirler çünkü katılacakları grubu seçme hakları yoktur. Eğer multimedya ile öğretimin etkili olup olmadığını bilmiyorsak, böyle olup olmadığını anlamak için en iyi ve en adil yöntem seçkisiz atamadır.

C Eğer sadece multimedya ile öğretim gören öğrenciler test edilmiş olsaydı, araştırmanın deseni bir tek grup ön test son test desen olurdu. Bu desende iç geçerliğe çok sayıda tehdit vardır; bu desene deney öncesi bir desen veya kötü bir deney denilmesinin nedeni budur. Bu araştırmada olası bir tehdit, test etkisidir; yani, öğrenciler çoğunlukla bir dersteki ilk testten ikincisine geçerken gelişme gösterirler çünkü test işlemlerine ve öğretmenin beklentilerine alışırlar. Bu gelişmenin, multimedya öğretiminin kullanılmadığı durumda bile ortaya çıkması beklenir. Bu araştırmadaki denk olmayan kontrol grup deseni böyle bir tehdidi ortadan kaldırır çünkü puanlardaki test etkisine bağlı herhangi bir artış, her iki grupta da ortaya çıkacaktır. Multimedya yöntemi uygulanan grupta, kontrol grubuna kıyasla, son testte gözlenecek *daha büyük* bir artış, multimedyanın etkisine bağlanabilir.

KISIM BEŞ

Araştırmanın Analizi ve Rapor Edilmesi

BÖLÜM ON BİR

Verilerin Analizi ve Analizin Yorumlanması: Kısım 1. Verilerin Betimlenmesi, Güven Aralıkları ve Korelasyon

BÖLÜMÜN ANA HATLARI

GENEL BAKIŞ
ANALİZİN ÖYKÜSÜ
BİLGİSAYAR TEMELLİ VERİ ANALİZİ
ÖRNEK: ORTALAMALARIN KARŞILAŞTIRILDIĞI BİR DENEY İÇİN VERİ ANALİZİ
Aşama 1: Verileri Tanımaya Başlama
Aşama 2: Verileri Özetleme
Aşama 3: Verilerin Anlamını Onaylamak İçin Güven Aralıklarının Kullanılması
ÖRNEK: KORELASYONEL ÇALIŞMA İÇİN VERİ ANALİZİ
Aşama 1: Verileri Tanımaya Başlama
Aşama 2: Verileri Özetleme
Aşama 3: Korelasyon İçin Güven Aralığının Hesaplanması
ÖZET

GENEL BAKIŞ

Veri analizinin başlıca amacı, gözlemlerin davranışa ilişkin bir iddiayı destekleyip desteklemediğini belirlemektir (Abelson, 1995). Bu iddia, madde bağımlısı annelerin çocuklarının, madde bağımlısı olmayan annelerden doğan çocuklara göre daha fazla öğrenme güçlüğü gösterdiği veya depresyonu engellemeyi amaçlayan bir programın işe yaradığı şeklinde olabilir. İddia ne olursa olsun çalışmamız, kanıtın niteliğine ve sunum şekline titizlikle dikkat edilerek hazırlanmalıdır. Niceliksel bir araştırma gerçekleştirildiğinde, kanıt öncelikli olarak topladığımız sayısal verilerdir. İkna edici bir sav öne sürebilmek için bu veriler içerisinde neyin aranacağını, bilginin nasıl özetleneceğini ve bilginin en iyi şekilde nasıl değerlendirileceğini bilmemiz gerekir.

Elbette veriler havadan gelmez; sonuçların belirli bir araştırma yöntemi (örneğin; gözlem, anket çalışması, deney) kullanılarak elde edildiğini varsayabiliriz. Eğer veri toplama aşamasında ciddi bir hata yapılırsa, veriyi "kurtarmak" için yapabileceğimiz hiçbir şey olmayabilir ve yapılacak olan en iyi şeyin baştan başlamak olması olasıdır. Bundan dolayı analiz edilecek verinin, hipotezin tanımlanmasına, bu hipotezin test edilmesi için uygun araştırma deseninin belirlenmesine, uygun tepki ölçümlerinin seçilmesine ve istatistiksel gücün değerlendirilmesine titizlikle dikkat edildikten sonra toplandığından emin olmalıyız. Ve elbette, talep özelliklerinin, deneyci yanlılıklarının, karıştırıcı değişkenlerin ya da araştırmanın diğer bozucu etkilerinin katkısını en aza indirecek biçimde verinin toplandığından emin olmak isteriz. Kısaca "iyi", yani iç ve dış geçerliğe sahip, duyarlılığı yüksek ve güvenilir bir araştırma verisini elde etmeyi amaçlarız.

Temel Kavramlar

Bilimsel ilkelere dayalı bir araştırmadan veri elde edildiğinden emin olduktan sonraki adımımız ne olmalıdır? **Veri analizinin** üç farklı fakat bağlantılı **aşaması** vardır: **Verileri tanımaya başlamak, verileri özetlemek** ve **verilerin anlamını onaylamak** (bk. Kutu 11.1). Yapılan çalışma ister gözlemsel bir çalışma olsun (bk. Bölüm 4), ister niceliksel veriye dayalı deneysel bir çalışma olsun (bk. Bölüm 6–8), veri analizinin ilk iki aşaması, yani verileri tanımaya başlamak ve verileri özetleme aşamaları hemen hemen aynı şekilde ilerlemektedir. Tarama araştırması (bk. Bölüm 5) veya iki değişkenin birlikte değişimine ilişkin kanıt arayan başka tür bir çalışma söz konusu olduğunda ise, verileri özetleme aşaması daha farklı biçimde ilerlemektedir. Veri analizinin aşamalarını göstermek için değişkenler arasındaki korelasyonu vurgulayanlar kadar, bir veya daha fazla grubun ortalama performansına odaklananları da içeren çeşitli araştırma örneklerini ele alacağız.

Veri analizinin üçüncü, yani verilerin anlamının onaylanması ile ilgili aşamasına dair farklı, fakat birbirini tamamlayıcı yaklaşımlar mevcuttur. Yaklaşımlardan birisi, evrendeğerlere ilişkin tahminin aralığı ve hassasiyeti ile ilgili kanıt sunmak üzere güven aralıklarından yararlanmaktadır. Diğer bir yaklaşım ise sıfır hipotezi anlamlılık testine (SHAT) dayanmaktadır. Bu iki yaklaşıma Bölüm 6'da kısaca değinilmişti ve daha önce söylediğimiz gibi bu iki yaklaşım birbiri ile bağlantılıdır; böyle olmakla birlikte, ikisi arasında önemli farklar da mevcuttur ve önce iki yaklaşıma ayrı ayrı değineceğiz, sonra her iki yaklaşımdan elde edilen bilginin analiz sürecinin sonunda nasıl birleştirilebileceğini göstereceğiz. Bu bölümde güven aralıklarını, Bölüm 12'de ise SHAT'ni tartışacağız. Buna ek

KUTU 11.1

VERİ ANALİZİNİN ÜÇ AŞAMASI

Veri analizinin üç temel aşaması, aşağıdaki şekilde betimlenebilir:

I **Verileri Tanımaya Başlamak** Bu ilk aşamada veriyi tanımak isteriz. Bu, keşfedici ve inceleyici bir aşamadır (Tukey, 1977). Bu aşamada veriyi dikkatlice inceler, veriyi hisseder ve hatta bazı araştırmacıların söylediği gibi, veri ile "arkadaş" oluruz (Hoaglin, Mosteller ve Tukey, 1991, s. 42). Sorduğumuz sorular arasında, "Bu sayı yığınında neler oluyor?", "Veride herhangi bir hata var mı?", "Veri bir anlam ifade ediyor mu yoksa veride bir 'bit yeniği' olduğundan şüphelenmeli miyiz?" (Abelson, 1995, s. 78) gibi sorular yer almaktadır. Bu aşamada sayıların dağılımının görsel olarak incelenmesi önemlidir. Başka bir deyişle, "Verinin görünümü neye benzemektedir?" sorusu sorulmalıdır. Verinin genel özelliklerine aşina olduğumuz, hataları kontrol ettiğimiz ve verinin bir anlam ifade ettiğinden emin olduğumuz noktada ancak bir sonraki aşamaya geçebiliriz.

II **Verileri Özetleme** İkinci aşamada amacımız veriyi anlamlı bir şekilde özetlemektir. Bu aşamada betimsel istatistikler ve grafiksel gösterimler önem kazanır. Yine bu aşamada, veri nasıl organize edilmelidir, veriyi tanımlamak ve özetlemek için en bilgilendirici yollar hangileridir, ilgilenilen faktörlerin bir fonksiyonu olarak bu çalışmada ne gerçekleşti, ne tür eğilimler ve örüntüler görmekteyiz, hangi grafiksel gösterim türü bu eğilim ve örüntüleri en iyi şekilde açığa çıkarabilir gibi sorular sorulur. Veri uygun şekilde özetlenebildiğinde, onaylama aşamasına geçmeye hazırız.

III **Verilerin Anlamının Onaylanması** Üçüncü aşamada verinin davranışla ilgili olarak ne söylediğine karar veririz. Veri, araştırmanın başlangıcında ortaya attığımız tahminî iddiamızı (araştırma hipotezini) desteklemekte midir? Eldeki kanıtlara dayanarak ne iddia edebiliriz? Bazen aradığımız şey, evet-hayır şeklinde bir kategorik yargıdır; bu durumda bir hükme varmak için yargıç veya jüri gibi davranırız. Hüküm vermeye yetecek kanıtımız var mı? Evet veya hayır: Bulunan etki gerçek mi? Bu aşamada, elde ettiğimiz sonucun sadece "şans eseri" ortaya çıktığını söyleyebilecek argümanlara karşı çıkabilmek için çeşitli istatistiksel tekniklerden yararlanabiliriz. Analizin bu aşamasında eğer uygunsa sıfır hipotezi testi yapılır. Bununla birlikte, yaptığımız değerlendirme bizi her zaman veriye ilişkin kategorik bir yargıya ulaştırmak zorunda değildir (örneğin, Schmidt, 1996). Başka bir deyişle, her zaman sonuçların "gerçekliğiyle" ilgili kesin bir ifadeyi amaçlamak zorunda değiliz. Davranışa ilişkin iddiamız, ilgilenilen değişkenle ilgili etki büyüklüklerinin olası aralıklarının değerlendirilmesine dayalı olabilir. Başka bir ifadeyle, belirli bir değişkenin varlığında gerçekleşmesi olası olan şey nedir? Bu türden bir değerlendirme yapabilmek için güven aralıkları bilhassa önerilmektedir (örneğin, Cohen, 1995; Hunter, 1997; Loftus 1996).

Onaylama süreci aslında veri analizinin keşfedici aşaması olan ilk aşamasında, verilerimizin neye benzediğine dair fikrimizin oluştuğu aşamada, başlar. Verinin genel özelliklerini incelerken, bulgumuzun ne olduğunu değerlendirmeye başlarız. Özetleme aşamasında, elde edilen gözlemlerin eğilimleri ve örüntüleri hakkında daha fazla şey öğreniriz. Bu durum, hipotezimizi onaylamak üzere geri bildirim sağlamaktadır. Veri analizindeki son aşama, verinin ne anlama geldiğine ilişkin bir karara vardığımızı vurgulamak için, onaylama aşaması olarak adlandırılmaktadır. Bununla birlikte, veri analizinin her aşamasında edinilen bilgi onaylama sürecine katkıda bulunmaktadır (örneğin, Tukey, 1977).

olarak Bölüm12'de, önemli bir kavram olan istatistiksel güç ve bunun güven aralıkları ve SHAT ile olan ilişkisini tartışacağız.

ANALİZİN ÖYKÜSÜ

- Veri analizi tamamlandığında, bulgularımızı açıklayan, alternatif yorumlara karşı çıkabilen ve vardığımız sonuçları meşrulaştıran, tutarlı bir metin yapılandırmamız gerekir.

Davranışla ilgili bir iddiaya ait ikna edici bir tartışma ortaya koymak, veriyi sadece analiz etmekten daha fazlasını gerektirir. İyi bir tartışma, iyi bir öykü gerektirir. Bir avukat, davayı kazanmak için jürinin dikkatini sadece davadaki olgulara çekmek yerine, aynı zamanda söz konusu olguları mantıklı ve tutarlı bir olay örgüsü içerisinde sunmalıdır. Eğer kanıtlar, suçlu olarak uşağı işaret ediyorsa, bu durumda suçu neden uşağın işlemiş olabileceğini (neden aşçının suçu işlemediğini) bilmek isteriz. Abelson (1995), araştırmalarda ortaya atılan tartışma ile ilgili olarak benzer bir noktaya değinmiştir:

> En üst düzeyde ikna edici bir etki sağlayacak bir istatistiksel argümana ulaşmak için, yüksek kalitede kanıtlar, somut ve iyi ifade edilmiş genel etkiler gereklidir ama yeterli değildir (s. 13).

Sonuç olarak, veri analizi tamamlandığında bulgularımızı açıklayan, alternatif açıklamaları eleyebilen ve vardığımız sonuçları meşrulaştıran tutarlı bir metin yapılandırmalıyız. Bölüm 12 ve 13'te uygun bir metin yapılandırmada size yardımcı olacak kılavuza değindiğimiz zaman, analiz öyküsüne geri döneceğiz.

BİLGİSAYAR TEMELLİ VERİ ANALİZİ

- Araştırmacılar genellikle verinin istatistiksel analizini gerçekleştirmek için bilgisayarları kullanırlar.
- Bilgisayar yazılımı kullanarak istatistiksel analiz yapabilmek için, araştırmacıların iyi düzeyde araştırma deseni ve istatistik bilgisine sahip olmaları gerekir.

Çoğu araştırmacının, verilerin istatistiksel analizini gerçekleştirmek için uygun yazılımı içeren bilgisayarlara hazır erişimi bulunmaktadır. İstatistiksel paket programı kullanarak bir analizi hazırlayabilmek ve gerçekleştirebilmek, ayrıca çıktıyı yorumlayabilmek, araştırmacıların öğrenmeleri gereken temel becerilerdir. Daha popüler olan paket programların bazıları BMDP, SAS, SPSS ve STATA gibi kısaltmalarla bilinmektedir. Muhtemelen sizin de bu yazılımların birine ya da daha fazlasına, psikoloji bölümünüzde veya kampüsünüzdeki bilgisayar laboratuvarında bulunan bilgisayarlarda, hatta kendi dizüstü bilgisayarınızda erişiminiz vardır.

Bilgisayar yazılımı kullanarak istatistiksel analiz yapabilmek, araştırmacının iyi düzeyde araştırma deseni bilgisine ve istatistik bilgisine sahip olmasını gerektirir. Bölüm 6, 7 ve 8'de çeşitli deneysel desenlere değinmiştik. Bu bölümlerdeki bilgiler, eğer bilgisayar temelli analiz yapmak istiyorsanız, gereklidir. Bir bilgisayar hangi araştırma desenini kullandığınızı ve bu desenin kullanılmasının altında yatan nedeni belirleyemez (bazı kullanıcı dostu programlar sizi yönlendirmek için hatırlatmalar sunuyor olsa bile). Bilgisayar temelli veri analizi yapmak için, kullanılan desenin türü (örneğin, seçkisiz gruplar veya tekrarlı ölçümler), bağımsız değişkenlerin sayısı (tek faktörlü veya çok faktörlü), her bağımsız değişkendeki düzey sayısı ve bağımlı değişken sayısı ile bunların her biri için ölçüm düzeyi gibi bilgileri girmeniz gerekir. Ayrıca araştırma hipotezlerinizi sözcüklere dökebiliyor olmanız ve araştırma hipotezlerinize uygun

istatistiksel testleri planlayabiliyor olmanız gerekir. Bir bilgisayar, hızlı ve etkin bir biçimde, betimsel ve çıkarımsal istatistikler için gerekli olan hesaplamaları yapabilir. Bununla birlikte, bilgisayarı bir araştırma aracı olarak etkili bir şekilde kullanabilmek için, hangi istatistiksel testi istediğinize ilişkin ve testte hangi verinin kullanılacağıyla ilgili belirli komutları bilgisayara vermelisiniz. Sonunda bilgisayar hesaplamaları tamamladığı zaman, analiz sonuçlarını gösteren çıktıyı doğru bir biçimde yorumlayabiliyor olmanız gerekir.

ÖRNEK: ORTALAMALARIN KARŞILAŞTIRILDIĞI BİR DENEY İÇİN VERİ ANALİZİ

Kaç sözcük biliyorsunuz? Başka bir deyişle, sözcük dağarcığınızın genişliği ne kadardır? Bu soruyu kendinize, CAT veya ACT gibi üniversite giriş sınavlarına hazırlanırken sormuş olabilirsiniz ya da LSAT veya GRE gibi profesyonel okul sınavlarına hazırlanmayı düşünürken aklınızdan geçmiş olabilir; bunun nedeni, tüm bu sınavların sözcük bilgisi üzerinde durmasıdır. Şaşırtıcı şekilde, bir kimsenin sözcük dağarcığı genişliğini tahmin etmeye çalışmak oldukça karmaşık bir görevdir (örneğin; Anglin, 1993; Miller ve Wakefield, 1993). Bu görevle ilgili sorunlar daha en başta kendini göstermektedir; örneğin, "sözcük" derken neyi kastettiğimizi düşündüğümüz zaman. "Oynamak, oynadı ve oynuyor" sözcükleri bir sözcük müdür yoksa üç farklı sözcük müdür? Altı heceli isimleri olan kimyasal bileşiklerin de dâhil olduğu son derece teknik veya bilimsel sözcükler de inceleme alanımıza giriyor mu? Peki uydurulmuş sözcükler ne olacak ya da köpeğinizin adı veya partnerinize hitap ederken kullandığınız sözcük? Oldukça doğrudan bir yöntem, kişinin İngilizce sözlükteki kelimelerin ne kadarını bildiğini sormaktır. Fakat burada bile zorluklarla karşılaşırız çünkü sözlükler genişlik ve kapsam açısından farklılaşmaktadır ve bu nedenle sonuçlar, sözcük örneklemini seçmek üzere belirlenen sözlüğe bağlı olarak değişecektir. Ve elbette sözcük bilgisi düzeyine ilişkin tahmin, bilginin nasıl test edildiğine bağlı olarak değişecektir. Çoktan seçmeli sorulardan oluşan testler, sözcüklerin tanımlarının yazılmasını gerektiren testlere göre bilgi düzeyini daha yüksek gösterecektir.

Kitabınızın yazarlarından birisi geçmişte sözcük dağarcığı genişliği konusunda çalışmıştır ve üniversite öğrencileri ile yaşlı yetişkinlerin sözcük dağarcıklarını inceleyen bir çalışma yapmıştır (bk. Zechmeister, Chronis, Cull, D'Anna ve Healy, 1995). Çalışmada tabakalı (harfe göre) seçkisiz örnekleme yöntemi kullanılarak bir İngilizce sözlükten 191 sözcük seçilmiştir. Sonra, beş şıklı çoktan seçmeli bir test hazırlanmıştır. Testte kelimenin doğru anlamını bulmayı zorlaştırmak için, bu kelimeye yakın anlamlı dört tuzak kelime de verilmiştir. Örneğin katılımcıların, "savaşçı" sözcüğünün anlamını, şu alternatifler arasından tanımaları istenmiştir: a. savaş, b. silah, c. yiğit, d. cesur, e. kahraman. Kelimeler üniversite çağındaki 26 öğrenciye (yaş ortalaması 18.5) ve 26 yaşlı yetişkine (yaş ortalaması 76) kitapçık hâlinde sunulmuştur. Önceki çalışmalar dikkate alındığında, yaşlı yetişkinlerden oluşan grubun gençlerden oluşan gruba göre sözcük bilgisi testinde daha yüksek puan almaları beklenmiştir.

Veri analizinin üç aşamasını gösterebilmek için bu araştırmanın verilerini kullanacağız.

Aşama 1: Verileri Tanımaya Başlama

- Veri analizine verinin genel özelliklerini incelemekle ve gerektiğinde veriyi düzenlemekle ya da "temizlemekle" başlarız.
- Eksik veri ve mümkün olmayan değer (örneğin, ölçek aralığının dışındaki sayılar) gibi hataları, ayrıca aşırı değerleri dikkatle kontrol etmek önemlidir.
- Gövde ve yaprak gösterimi, özellikle veri kümesinin genel özelliklerinin görselleştirilmesi ve aşırı değerlerin saptanması için kullanışlıdır.
- Veriler sayısal, görsel ve sözel olarak etkili bir biçimde özetlenebilir; verilerin iyi bir şekilde tanımlanabilmesi için, sıklıkla bu üç özetleme şekli de kullanılır.

Verinin Temizlenmesi Verinin genel özelliklerini incelemekle, gerektiğinde veriyi düzenlemek ve "temizlemekle" işe başlamak isteriz (Mosteller ve Hoaglin, 1991). Eksik veri ve mümkün olmayan değer (örneğin, ölçek aralığının dışındaki sayılar) gibi hataları dikkatlice kontrol ederiz. Hatalar, katılımcıların ölçeği yanlış kullanmasından (örneğin, önem sırasını ters çevirmek) veya veriyi bilgisayara giren kişinin bir sayıyı atlamasından ya da bir rakamın sırasını değiştirmesinden kaynaklanabilmektedir. Bir metni yazarken çoğumuz, yazım hatalarını yakalamak için "yazım denetimcisine" güveniriz. Ne yazık ki bilgisayara girilen sayısal hataları saptayabilecek böyle bir sistem yoktur (fakat bk. Kaschak ve Moore, 2000). Analiz sürecine devam etmeden önce verinin temizlenmiş olup olmadığına, araştırmacı karar verecektir.

Anormalliklerin ve hataların saptanması çok önemlidir. Daha önceden görmüş olduğumuz gibi, bir anormallik bazen veri kaydındaki bir hataya işaret eder; örneğin, katılımcıların 7'li bir ölçekteki yanıtları arasında 8 rakamına rastlanması veya üniversite öğrencilerinden oluşan bir örneklemde IQ puanının 10 olarak kaydedilmesi gibi. Diğer anormallikler aşırı değerlerle ilgilidir. Bir aşırı değer, bir dizinin uç kısmında bulunan sayıdır ve olası değerlerin içinde yer alsa da, verinin genel eğilimine uymuyormuş izlenimi verir. Örneğin tepki süresinin kullanıldığı bir çalışmada çoğu tepkinin 1500 ms'nin altında olmasını bekleriz ve 4000 ms'lik bir tepki zamanı ile karşılaştığımızda şaşırabiliriz. Geniş bir veri kümesinde, diğer tüm değerlerin neredeyse hepsi 1500'ün altında ise, 4000 değeri kesinlikle aşırı değer olarak değerlendirilecektir. Tepki zamanı çalışmalarında bu tür değerler, örneğin, katılımcı hapşırdığında, dalgınlıkla ekrandan başka bir yere baktığında ya da yanlışlıkla veri toplama sürecinin sona erdiğini zannedip gitmeye hazırlandığında ortaya çıkabilir. Bir anketi yanıtlayan bir katılımcı sorulardan birini yanlış okuyabilir veya veri kümesindeki herhangi bir yanıttan daha da uç noktada bulunan bir yanıt verebilir. Ne yazık ki araştırmacılar aşırı değer kavramının bir tek tanımını kullanmazlar; birkaç farklı, "göz kararı" olarak nitelenebilecek ölçüt kullanılır (bk. örneğin, Zechmeister ve Posavac, 2003).

Bir veri kümesinde anormalliklere rastlandığında, bunların sonraki analizlerden çıkarılıp çıkarılmayacağına karar germemiz gerekir. Hata olduğuna kesin olarak karar verilen anormallikler düzeltilmelidir veya veri kümesinden çıkartılmalıdır; fakat bunu yaparken araştırmacı, veri analizinden çıkardıklarını araştırma raporunda belirtmeli ve eğer mümkünse anormalliğin kaynağını da açıklamalıdır.

Veri analizinin ilk aşamasında, puanların dağılımını anlamlı bir biçimde açıklamanın yollarını ararız. Verideki değişkenlik neye benzemektedir? Veriler normalden kaymış mı yoksa görece normal mi dağılmış? Analizin ilk aşamasının amaçlarından birisi, analize devam etmeden önce verilerin dönüştürülmesinin gerekip gerekmediğini belirlemektir. Verinin dönüştürülmesi "yeniden ifade etme" sürecidir (Hoaglin, Mosteller ve Tukey, 1983). Nispeten basit dönüştürmelere örnek olarak inç ölçümünü fit olarak, fahrenhaytı selsiyus olarak veya "doğru sayısı"nı "doğru yüzdesi" olarak ifade etmek gösterilebilir. Bazen karmaşık istatistiksel dönüştürmeler de kullanışlıdır.

Veri kümesine dair ilk izlenimi edinmenin en iyi yolu, veriyi resmetmektir. Bilgisayar temelli veri analizinin bir avantajı, verinin resminin nasıl değiştiğini görebilmek için, çeşitli gösterim seçeneklerini kullanarak (örneğin; frekans poligonları, histogramlar) hızlı ve kolay bir şekilde veriyi grafik hâlinde gösterebiliyor olmamız ve aynı kolaylıkla ölçeği de değiştirebiliyor olmamızdır (örneğin, inç ölçümünü fite çevirmek). En basit şekliyle, veri kümemizi görselleştirmenin farklı yollarını kullanarak verimizi daha yakından tanımış oluruz. Hangi görsel sunum türü verimiz hakkında en çok bilgi ortaya çıkarmaktadır? Farklı şekilde tanımlanmış eksenler ile çizilmiş grafikleri karşılaştırarak, verimize ilişkin ne öğrenmiş oluruz? Poligon mu, histogram mı daha çok bilgi verir? Bir resim, sadece 1000 kelimeye bedel değildir, aynı zamanda resim, 1000 sayıyı hızlıca özetleyebilir. Verimize ait farklı resmetme türlerine aşina oldukça, bazı resimlerin diğerlerinden daha iyi olduğunu öğreniriz.

Sözcük bilgisi çalışması örneğimizdeki veri, 191 olası anlam arasında doğru şekilde tanımlanan sözcük anlamlarının sayısını göstermektedir. Çoktan seçmeli testlerde katılımcılar doğru yanıtı bilmiyor olsalar bile, yanıt şans eseri doğru olabileceği için, yanıtların tahmin edilmesinin yarattığı bozucu etkiyi azaltmak üzere, katılımcıların tepkileri üzerinde standart bir düzeltme işlemi uygulanmıştır. Ayrıca yaşlı yetişkin grubuna verilen kitapçıklarda iki basım hatasına rastlanmıştır; bu nedenle söz konusu maddeler, yapılacak olan analizlere dâhil edilmemiştir. Ayrıca, test kitapçıkları üzerinde yapılan incelemede, bazı yaşlı katılımcıların sayfa atladığı fark edilmiştir. Bu nedenle, söz konusu katılımcılar için olası sözcük miktarı düşürülmüştür. Bu gibi sorunlar nedeniyle toplam olası tepki miktarının katılımcıdan katılımcıya değiştiğini hesaba katabilmek için veri, doğru tepki yüzdesini gösterecek şekilde dönüştürülmüştür.

Veri kümesinin temizlenmesinin ardından, analizin ilk aşamasında araştırmacılar aşağıdaki verileri elde etmişlerdir. Bu veriler doğru yüzdesi cinsinden, üniversite öğrencilerinin ve yaşlı yetişkinlerin çoktan seçmeli test performansını ifade etmektedir.

Üniversite öğrencileri (n = 26): 59, 31, 47, 43, 54, 42, 38, 44, 48, 57, 42, 48, 30, 41, 59, 23, 62, 27, 53, 51, 39, 38, 50, 58, 56, 45.

Yaşlı yetişkinler (n = 26): 70, 59, 68, 68, 57, 66, 78, 78, 64, 43, 53, 83, 74, 69, 59, 44, 73, 65, 32, 60, 54, 64, 82, 62, 62, 78.

Temel Kavram

Gövde ve Yaprak Gösterimleri **Gövde ve yaprak gösterimi**, özellikle verinin genel özelliklerinin görselleştirilmesinde ve aşırı değerlerin saptanmasında kullanışlıdır (Tukey, 1977). Gövde ve yaprak gösterimi adı, sayısal dizimde ilk basamakların "gövdeler" ve sonraki basamakların "yapraklar" biçiminde geleneksel kullanımından gelmektedir.

Üniversite öğrencilerinin yukarıda söz edilen sözcük dağarcığı araştırmasında elde ettikleri verilere ait gövde ve yaprak gösterimi aşağıdadır:

2*	3
2	7
3*	01
3	889
4*	12234
4	5788
5*	0134
5	67899
6*	2

Buradaki ilk basamağı gösteren rakamlar onlar basamağıdır (örneğin, 2-, 3-, 4-), sonraki basamaklar ise ilk basamak veya en önemli basamakları takip eden rakamlardır; bu örnekte sonraki basamaklar birler basamağıdır (örneğin,, -5, -6, -8). Bu gösterimde ilk basamaklar dikey olarak, en üst sırada en küçük sayı yer alacak şekilde düzenlenmiştir. İlk basamağı ise, artan bir sırada olmak üzere, dağılımda görülebildiği miktar kadar sonraki basamak takip etmektedir. Gösterimdeki her satır, sonrasında gelen yapraklarıyla birlikte, bir gövdeyi belirtmektedir (Tukey, 1977). Örneğin yukarıdaki gösterimde 3 sayısıyla gösterilen gövdenin üç yaprağı bulunmaktadır, bunlar 8, 8, 9'dur ve dağılımdaki 38, 38 ve 39 sayılarına karşılık gelmektedir. Eğer çok miktarda rakam gösteriliyorsa veya tüm veri seti sadece birkaç ilk basamağı gösteren rakam içeriyorsa, ilk basamaklar sayı aralığının ilk yarısını göstermek amacıyla geleneksel olarak yıldız (*) ile işaretlenmektedir (bk. Tukey, 1977). Örneğin 5*, 0, 1, 2, 3 ve 4 yapraklarının gövdesidir (50-54 aralığındaki sayılar); diğer bir ilk basamak olan 5 sayısı ise (* işareti olmadan), 5, 6, 7, 8, 9 yapraklarının gövdesidir (55-59 aralığındaki sayılar). Yukarıdaki gösterimde örneğin 2* gövdesinin bir yaprağı vardır (3), benzer şekilde 2 gövdesinin de bir yaprağı vardır (7) ve sırasıyla 23 ve 27 sayılarını göstermektedirler.

Veride ayrıca birden fazla ilk basamak bulunabilir. Örneğin puanlar 50 ve 150 arasında değişiyorsa tek rakamdan oluşan ilk basamaklar 100'den küçük sayılar için (8-, 9-, vs.), iki rakamdan oluşan ilk basamaklar ise 100 ve 100'den büyük sayılar için (10-, 11-, 12, vs.) kullanılır.

Gövde ve yaprak gösterimi, dağılımdaki her veri noktasını tanımlayabilmemizi sağlar. Dahası, bu gösterimin önemli bir avantajı, dağılımın şeklini ve eğer varsa aşırı değerleri açık bir biçimde göstermesidir.

Yirmi altı üniversite öğrencisinin sözcük dağarcığı verilerine ait gövde ve yaprak gösterimini dikkatlice inceleyin. Ne görüyorsunuz? Dağılımın genel biçimi "normal" mi (simetrik ve çan şeklinde dağılım) yoksa eğik mi (asimetrik ve puanların bir yöne doğru kaydığı dağılım)? Puanlar yayılmış mı, yoksa belirli bir değer etrafında toplanma eğilimi mi gösteriyor? Normal olmayan değerler var mı? Bu verilere ait gövde ve yaprak gösteriminin, değerlerin yüzde 40 ve 50'ler civarında toplandığını ve dağılımın negatif yönde biraz çarpık olduğunu ("kuyrukların", dağılımın düşük değerleri içeren kısmına, negatif kısmına doğru nasıl kaydığına dikkat ediniz) gösterdiği söylenebilir. Herhangi bir aşırı değer görünmemektedir (örneğin, tek basamaklı yüzde veya 60'lı değerlerin ötesinde bir yüzde yoktur).

Karşılıklı olarak iki gövde ve yaprak gösterimini kullanmak, özellikle iki grup verinin karşılaştırılması için aydınlatıcı olabilir. Sonraki sayfada yer alan

gösterimi ele alın. Sağdan sola artan bir dağılım içinde sonraki basamakların (örneğin, 997 5), soldan sağa artan bir dağılım biçiminde de diğer yaprakların (5 67899) gösterildiği ortak gövdeler kullanılmıştır. Bu, ilk dağılımda 57, 59 ve 59 puanlarının yer aldığı, ikinci dağılımda ise 56, 57, 58, 59 ve 59 puanlarının bulunduğu anlamına gelmektedir. Karşılıklı gövde ve yaprak gösterimleri eğer araştırmacı sosyoekonomik düzey, yaş, cinsiyet veya diğer başka önemli bir özellik açısından farklılaşan iki grubu karşılaştırıyorsa, örneğin belirli bir anket sorusuna verilen yanıtların karşılaştırılmasında olduğu gibi, uygun biçimde kullanılabilir.

Sözcük dağarcığı araştırmasının iki koşulu için karşılıklı gövde ve yaprak gösterimleri şuna benzeyecektir:

Yaşlı Yetişkinler		Üniversite Öğrencileri
	2*	3
	2	7
2	3*	01
	3	889
43	4*	12234
	4	5788
43	5*	0134
997	5	67899
44220	6*	2
98865	6	
430	7*	
888	7	
32	8*	

Soldaki, yaşlı katılımcılar için olan gösterime bakınız. Bunu nasıl yorumlarsınız? Verilerde uçta bir değer, bir aşırı değer bulunmakla birlikte, oldukça normal dağılmış gibi görünmektedir. "32" değeri verinin geri kalanına ait görünmüyor (dağılımın merkezine olan uzaklıklarına göre aşırı değerleri tanımlamanın yolları vardır ve bazı bilgisayar programları bunu otomatik olarak yapmaktadır). Katılımcının özelliklerine ilişkin ek bir bilgi olmadan (örneğin, olası görme sorunları), deneyciler bu veriyi analizden çıkarmak için herhangi bir neden olmadığını düşünebilirler. Bu olası aşırı değerin varlığı, bu gruptaki değişkenliği, olmadığı duruma kıyasla, kaçınılmaz olarak artıracaktır. Yine de bazı veri kümelerinin diğerlerine göre daha fazla değişkenlik gösterdiğini kabul etmeliyiz. Örneğin, bu çalışmadaki yaşlı yetişkinler basitçe, üniversite öğrencilerinden oluşan örnekleme göre daha heterojen bir grubu temsil ediyor olabilirler. Bu noktada uygulanması gereken bir ilke vardır: Veri toplama aşamasında katılımcılarınızdan toplayabildiğiniz kadar çok bilgi toplayın. Bir uç değerin hata sonucu veya çalışma ile ilgili olmayan nedenlerden ötürü oluştuğunu belirlemediğiniz sürece, söz konusu değer gerçek puan olarak kabul edilmelidir.

Karşılıklı gövde ve yaprak gösterimlerinin her iki dağılıma ilişkin olarak ne gösterdiğine bakınız. Bir dereceye kadar iki grubun puanlarının örtüştüğünü hemen görebilmelisiniz; bununla birlikte yaşlıların yer aldığı grupta, üniversite öğrencilerinin bulunduğu gruba göre çok daha fazla 60'ın üzerinde puan var. Veriye ait bu "resim", yaşlı yetişkinlerin, sözcük dağarcığı testinde üniversite öğrencilerine göre daha yüksek bir performans gösterdiği düşüncesini onaylamaya adaydır.

Sonuç: Veri analizinin ilk aşamasında –verimizi tanımaya başlama sürecinde– şunları belirlemeliyiz:

(a) Veri kümesindeki hataların niteliği ve sıklığı; hatalar var ise düzeltmelerin yapılıp yapılamayacağı veya verinin çıkarılmasının gerekli olup olmadığı;
(b) Aşırı değerlerin de dâhil olduğu, normal olmayan değerlerin varlığı; var ise bu değerlerin ortaya çıkmasının nedenleri; bu değerler ile ilgili olarak ne yapılabileceği (olduğu gibi bırakmak veya çıkarmak);
(c) Sayısal verilerin dağılımının genel özellikleri ve şekli ve
(d) Veriyi daha iyi ifade etmek için kullanılabilecek alternatif yöntemler.

Aşama 2: Verileri Özetleme

- Merkezî eğilim ölçüleri, ortalama, ortanca ve modu içerir.
- Yayılımın ya da değişkenliğin önemli ölçüleri aralık ve standart sapmadır.
- Ortalamanın standart hatası, ortalamaların teorik örnekleme dağılımının standart sapmasıdır ve evrenin ortalamasını ne kadar iyi tahmin edebildiğimizin ölçüsüdür.
- Etki büyüklüğü ölçüleri, bağımsız değişken ve bağımlı değişken arasındaki ilişkinin gücünü, örneklemin büyüklüğünden bağımsız olarak gösterdiği için önemlidir.
- İki ortalamanın karşılaştırılmasında kullanılan önemli bir etki büyüklüğü ölçüsü Cohen *d*'dir.

Veriler sayısal, görsel ve sözel olarak etkili bir biçimde özetlenebilir. Verinin iyi şekilde tanımlanmış hâli sıklıkla bu üç özetleme türünü de kullanır. Bu bölümde bazı grafikler gösterecek olsak da, daha çok verileri sayısal olarak özetleme yollarına, yani betimsel istatistiklerin kullanımına odaklanacağız. Verileri özetlemek amacıyla grafik çizmeye ilişkin bilgiler Bölüm 13'te ayrıca yer almaktadır. Verilerin sözel olarak tanımlanması da Bölüm 13'ün başlıca konularındandır (özellikle, araştırma raporunun bulgular kısmının yazımına ilişkin önerilere bakınız).

Sözcük dağarcığı çalışmasına ait veriler merkezî eğilim, yayılım, ortalamanın standart hatası ve etki büyüklüğü ölçüleri kullanılarak özetlenecektir.

Temel Kavramlar

Merkezî eğilim Merkezî eğilim ölçüleri ortalama, ortanca ve modu içerir. **Merkezî eğilim ölçüleri** aslında isimlerinin gerektirdiği şeyi yapmaktadır: Ölçümlerin, hangi puan etrafında toplanma eğilimi olduğunu gösterir. Bunlardan **Mod**, en kaba merkezî eğilim ölçümüdür: Basitçe, frekans dağılımında en sık gözlenen puanı gösterir. Eğer dağılımdaki iki puan, diğer puanlara göre daha yüksek bir sıklıkla gözlenmişse ve bu iki puan, frekans dağılımında farklı bölgelerde yer alıyorsa, söz konusu dağılım, bimodal adını almaktadır (iki moda sahip anlamında).

Temel Kavram

Ortanca (Medyan), frekans dağılımının orta noktasıdır. Tüm puanlar en küçükten en büyüğe sıralandıktan sonra, bu dağılımı iki eşit parçaya bölen değer ortanca olur. Şu veri kümesini ele alın: 4, 5, 6, 7, 8, 8. Bu veriler için ortanca 6.5'tir. Dağılımdaki puanların sayısı çift ise, ortanca, ortadaki iki değerin ortalaması alınarak hesaplanmaktadır [buradaki örnekte, (6 + 7) / 2 = 6.5].

Sıralı puanların sayısı tek ise ortanca, en ortadaki değerdir. Şu sayı grubu için 4, 5, 6, 17, 18 ortanca 6'dır. En yüksek değer 180 olsaydı, ortancanın 180 değil, yine 6 olacağına dikkat ediniz. Dağılımda uç değerler yer alıyorsa, merkezî eğilimin en iyi ölçüsü ortancadır çünkü ortanca, uç değerlerden, ortalamaya kıyasla, daha az etkilenmektedir.

Temel Kavram

Ortalama, en sık rapor edilen merkezî eğilim ölçüsüdür ve puanların toplamının, bu toplama katılan puan sayısına bölünmesi ile bulunur. Evren ortalaması μ ile (Yunan alfabesinde mu harfi) gösterilir. Örneklem ortalaması metinde rapor edilirken, örneğin bulgular bölümünde, ort. ile gösterilir ($\bar{X}$ sembolü ise ["X çizgi" olarak okunur] istatistiksel formüllerde kullanılmaktadır). Dağılımda uç değerler bulunmadığı sürece, merkezî eğilimin ölçüsü olarak her zaman ortalama rapor edilmelidir. "Ortalama" puan kavramından bahseden bir kimse genellikle aritmetik ortalamayı kastetmektedir. Sözcük bilgisi çalışmasındaki grupların merkezî eğilim ölçüleri aşağıdaki gibidir:

	Üniversite	Yaşlı Yetişkin
Ortalama (ort.)	45.58	64.04
Ortanca	46.00	64.50
Mod	38,42,48,59	78

Görüldüğü gibi üniversite öğrencilerinin ortalama performansı, yaşlı yetişkinlerin ortalama performansından çok daha düşüktür. Bu durum, karşılıklı gövde ve yaprak gösteriminde gördüğümüz şeyi doğrulamaktadır: Yaşlılardan oluşan grup, genel olarak üniversite öğrencilerinden daha yüksek bir performans göstermiştir. Her iki grupta da ortalama ve ortancanın benzer oldukları görülmektedir; bu nedenle, gövde ve yaprak gösterimine baktığımızda yaşlılardan oluşan örneklemde uç değere rastlamış olsak bile bu değer, merkezî eğilim ölçüsü olan ortalamadan pek fazla uzaklaşmış gibi görünmemektedir. Üniversite öğrencilerinin verisinde her biri iki kez gözlenen, birden fazla mod yer almaktadır; yaşlılardan oluşan grupta ise en sık rastlanan puan 78'dir ve sadece üç kez gözlenmiştir. Görülebildiği gibi mod, bu küçük veri kümelerini özetlemede pek faydalı olamamıştır.

Yayılım ya da Değişkenlik Merkezî eğilim ölçüsü rapor ettiğiniz zaman, ardından uygun bir **yayılım ölçüsüne (değişkenlik)** değinmelisiniz. Merkezî eğilim ölçüleri, frekans dağılımındaki puanların, etrafında toplanma eğilimi gösterdiği "merkez" puanı gösterir; yayılım ölçüleri ise dağılımın genişliğini ya da değişkenliğini gösterir.

Temel Kavramlar

Yayılımın en kaba ölçümü (merkezî eğilim ölçülerindeki modda olduğu gibi) **aralıktır.** Aralık, dağılımdaki en düşük ve en yüksek puan ile gösterilmektedir. Örneğin 1, 3, 5, 7 puanlarından oluşan küçük dağılımın aralığı 1–7'dir.

En sık kullanılan yayılım ölçüsü, (merkezî eğilim ölçülerindeki ortalamada olduğu gibi) **standart sapmadır.** Standart sapma, puanların ortalamadan genelde ne kadar uzaklaştığını göstermektedir. Standart sapma, puanların,

ortalamadan farklarının kareleri toplamının puan sayısına bölümünün karekökünе eşittir.

Şu anda konumuz dışında olan bazı nedenlerle, ortalamadan sapmaların karelerinin ortalaması hesaplanırken, N yerine N–1'e bölme işlemi kullanılmaktadır, böylece evrenin standart sapmasının yansız bir tahmini örneklemden hareketle yapılabilmektedir. Evrenin standart sapması σ ile (Yunan alfabesindeki sigma harfi) sembolize edilmektedir; örneklemdeki puanların standart sapması metin içerisinde S ile gösterilmektedir fakat istatistiksel formüllerde sıklıkla s ile sembolize edilmektedir. Çeşitli çıkarımsal istatistiklerin hesaplanması için önemli bir yayılım ölçüsü olan varyans ise standart sapmanın karesidir ve s^2 ile gösterilmektedir.

Sözcük bilgisi çalışmasındaki iki grubun değişkenlik ölçüleri aşağıdadır:

	Üniversite	Yaşlı Yetişkin
Aralık	23–62	32–83
Varyans (s^2)	109.45	150.44
Standart sapma (S)	10.46	12.27

Dikkat ederseniz, gövde ve yaprak gösterimi, yaşlı yetişkinler arasında daha geniş bir yayılım olduğunu göstermişti; S ile, dağılımın bu özelliğini sayısal olarak göstermiş olduk.

Ortalamanın Standart Hatası Çıkarımsal istatistikler gerçekleştirilirken, evrene ait ortalamayı (μ) tahmin etmek (veya çıkarsamak) için örneklemin ortalamasını ($\bar{X}$) kullanırız. Çoğu zaman, $\bar{X}$'ten hareketle μ tahmin edilirken ne büyüklükte bir hata yapıldığını belirleyebilmek yararlıdır. Matematikteki merkezî limit teoremi, sonsuz sayıda ve eşit büyüklükte örneklem seçtiğimizde ve bu örneklemlerin her birisi için $\bar{X}$ değerini hesapladığımızda, bu örneklem ortalamalarının ortalamasının ($\mu_{\bar{X}}$) evren ortalamasına (μ) eşit olacağını söylemektedir ve örneklem ortalamalarının standart sapması da evrenin standart sapmasının ($\mu_{\bar{X}}$), örneklem büyüklüğünün (N) karekökünе bölümüne eşit olacaktır. Ortalamaya ait bu teorik örnekleme dağılımına **ortalamanın standart hatası** ($\sigma_{\bar{X}}$) adı verilmektedir ve aşağıdaki şekilde tanımlanmaktadır:

Temel Kavram

$$\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{N}}$$

Tipik olarak, evrenin standart sapmasının ne olduğunu bilemeyiz, bu nedenle örneklemin standart sapmasını (s) kullanarak bu değeri tahmin ederiz. Sonrasında, aşağıdaki formülü kullanarak **ortalamanın tahminî standart hatasını** hesaplayabiliriz:

Temel Kavram

$$s_{\bar{X}} = \frac{s}{\sqrt{N}}$$

$s_{\bar{X}}$'in küçük bir değere sahip olması evren ortalamasına ait iyi bir tahmin yaptığımız anlamına gelir; bununla birlikte, büyük değere sahip bir $s_{\bar{X}}$, ortalamaya ilişkin yalnızca kaba bir tahminimiz olduğunu gösterir. Ortalamanın standart hatasına ait formül, örneklemden hareketle evren ortalamasını tahmin etme becerimizin, örneklem büyüklüğüne (geniş örneklemler daha iyi tahmin yapmayı sağlamaktadır) ve örneklemin standart sapması tarafından tahmin edilebildiği kadarı ile, evrenin gösterdiği değişkenliğe (evrendeki puanlar daha

az değişkenlik gösterdikçe, evren ortalaması tahminimiz daha iyi olmaktadır) bağlı olduğunu göstermektedir. Daha sonra göstereceğimiz gibi, ortalamanın standart hatası, güven aralıklarının belirlenmesinde önemli bir rol oynar ve sıklıkla araştırmanın bulgularını özetleyen şekillerde örneklem ortalaması ile birlikte gösterilir.

Etki Büyüklüğü Ölçüleri Deney yapılırken bağımsız değişkenin bir etkisinin olup olmadığını ve eğer etkisi var ise, bu etkinin büyüklüğünü belirlemek isteriz. Bölüm 6'da etki büyüklüğü kavramına değinilmişti. Etki büyüklüğü ölçüleri ya da daha genel kullanımıyla "etki genişliği" ölçüleri (bk. Kirk, 1996), bağımsız ve bağımlı değişken arasındaki ilişkinin gücü hakkında örneklem büyüklüğünden bağımsız olarak bilgi verdikleri için önemlidir (özellikle bk. Grissom ve Kim, 2005).

Deneysel araştırmalarda iki ortalama karşılaştırıldığı zaman sık kullanılan bir etki büyüklüğü ölçüsü Cohen *d*'dir. Bu, bağımsız değişkenin düzeyleri için, ortalamaların grup içi standart sapmalara bölünmesiyle elde edilen değerler arasındaki farkı ölçen bir orandır. Standart sapmanın, yaklaşık olarak puanların grup ortalamasından genelde ne kadar uzaklaştığını göstermekte olduğunu hatırlayınız. Standart sapma, puanların, ortalamanın etrafına "dağılımının" bir ölçüsüdür ve grup içi standart sapma söz konusu olduğunda bize bireysel farklılıklardan kaynaklanan "hatanın" miktarını (yani bireylerin tepkilerinin nasıl farklılaştığını) göstermektedir. Standart sapma, ortalamalar arasındaki farkı değerlendirmek için kullanışlı bir ölçü olarak işlev görmektedir. Şöyle ki, bir deneyde bağımsız değişkenin etkisinin "büyüklüğü" (bağımsız değişken için, grup ortalamaları arasındaki fark) her zaman puanların yayılış derecesinin ortalaması olarak ifade edilmektedir.

Etki büyüklüğü ölçüsü olan *d*, merhum istatistikçi Jacob Cohen (*d*'ye ilişkin daha fazla bilgi için bakınız, Cohen, 1988) tarafından örneklem ortalamaları arasındaki farkın, ortak evrenin standart sapmasına bölümü olarak tanımlanmıştır.

$$\text{Cohen } d = \frac{\overline{X}_1 - \overline{X}_2}{\sigma}$$

Evrenin standart sapması (σ), grupların grup içi değişkenliklerinin birleştirilmesiyle ve bunun her iki grupta toplam puan sayısına (N) bölünmesiyle elde edilmektedir. Ortak evrenin standart sapması için olan ve örneklem varyanslarının kullanıldığı bir formül aşağıdaki gibidir:

$$\sigma = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{N}}$$

Burada:

n_1 = Grup 1'in örneklem büyüklüğü
n_2 = Grup 2'nin örneklem büyüklüğü
s_1^2 = Grup 1'in varyansı
s_2^2 =Grup 2'nin varyansı
$N = n_1 + n_2$

Eğer çok fazla grup içi değişkenlik söz konusu ise (yani grup içi standart sapma geniş ise), *d* için bölen büyük olacaktır. Böyle geniş bir grup içi değişkenlik söz konusuyken, bağımsız değişkenin etkisini gözleyebilmek için, iki grup ortalaması arasındaki farkın büyük olması gerekir. Grup içi değişkenlik az olduğunda ise (*d* için bölen küçük olduğunda), ortalamalar arasındaki benzer bir farklılık, daha geniş bir etki büyüklüğüne karşılık gelir. Etki büyüklükleri standart sapma cinsinden ifade edildiği için, farklı bağımlı değişkenlerin kullanıldığı deneylere ait etki büyüklüklerinin karşılaştırılmasında kullanılabilirler. Örneğin, üniversite öğrencilerinin ve yaşlıların sözcüklerin anlamını ayırt ettikleri araştırmanın etki büyüklüğü ile sözcük tanımlarını anımsama performansı açısından benzer grupları karşılaştıran başka bir çalışmanın etki büyüklüğü doğrudan karşılaştırılabilir. Bu türden karşılaştırmalar, çok sayıda çalışmayı inceleyerek belirli bir bağımsız değişkenin etkisini özetleyen meta-analizlere temel teşkil eder (bk. Bölüm 6).

d oranlarını yorumlamamızı kolaylaştıran bazı kılavuzlar mevcuttur. J. Cohen (1992), etki büyüklüklerine ilişkin olarak, küçük, orta ve büyük olmak üzere kullanışlı bir sınıflama hazırlamıştır. Her etki büyüklüğü sınıfı nicel olarak ifade edilebilmektedir; örneğin, iki grubun karşılaştırıldığı bir deneyde orta düzeydeki etki için *d* değeri .50'dir; küçük ve büyük etkiler için *d* değerleri sırasıyla .20 ve .80'dir. Etki büyüklüklerinin bu türden ifadesi, özellikle benzer çalışmalara ait bulguların karşılaştırılmasında kullanışlıdır.

Araştırmacıların, ortalamalar arasındaki standardize edilmiş farkları, biraz farklı şekilde tanımladıklarını da belirtelim. (örneğin, bk. Cohen, 1988; Kirk, 1996; Rosenthal, 1991). Hangi etki büyüklüğü yönteminin kullanılacağı araştırmacıya kalmış bir şeydir. Psikoloji literatüründe ölçümler arasındaki farklılıkları düşündüğümüzde, bir araştırma raporunda etki büyüklüğünün nasıl hesaplandığının ayrıntılı bir şekilde tanımlanması çok önemli hâle gelir.

Sözcük bilgisi çalışması için Coden *d* kullanılarak hesaplanan etki büyüklüğü aşağıdadır:

$$d = \frac{\bar{X}_1 - \bar{X}_2}{\sigma} = \frac{64.04 - 45.58}{\sqrt{\frac{(26-1)(150.04) + (26-1)(109.45)}{52}}} = 1.65$$

Buradaki 1.65 değerini yorumlayabilmek için, J. Cohen'in (1992), küçük etki büyüklüğü için *d* = .20, orta etki büyüklüğü için *d* = .50 ve geniş etki büyüklüğü için *d* = .80 biçimindeki sınıflamasını kullanabiliriz. Buradaki değerimiz .80'den büyük olduğu için, "yaş" değişkeninin sözcük bilgisi üzerinde geniş bir etkisinin olduğu sonucuna varabiliriz.

Sonuç Veri analizinin ikinci, yani özetleme aşamasında şunları belirlemeliyiz:

(a) çalışmadaki her koşul ya da grup için merkezî eğilim (örneğin, ortalama);
(b) çalışmanın her koşulu için standart sapma gibi yayılım (değişkenlik) ölçüleri;
(c) her ana bağımsız değişken için etki büyüklüğü ve

(d) verinin grafiksel olarak en iyi şekilde nasıl özetlenebileceği (örneğin, farklı koşullara ilişkin ortalama performansı gösteren şekil).

Not: Her ne kadar sözcük bilgisi çalışmasındaki iki gruba ait ortalama performansı gösteren bir grafik çizilebilirse de, genellikle sadece iki gruba ait ortalama söz konusu olduğunda grafiğe gerek yoktur. Grafiksel özetlemeler, iki gruptan daha fazla grubun yer aldığı çalışmalar için olan bulguları özetlerken daha önemlidir.

Aşama 3: Verilerin Anlamını Onaylamak İçin Güven Aralıklarının Kullanılması

- Verilerin bize ne söylediğini onaylamak için kullanılan önemli bir yaklaşım, ortalama veya iki ortalama arasındaki fark gibi evrendeğerler için güven aralıklarının belirlenmesidir.

Temel Kavram

Veri analizinin üçüncü aşamasında, veriye aşina olduğumuz durumda, verinin özetlenmesine ilişkin ölçümlerden elde edilen kanıtlara ilişkin izlenimimizi onaylamaya çalışırız. Bu üçüncü aşamadaki temel yaklaşımlardan biri, **evrendeğer güven aralığının** hesaplanmasıdır. Güven aralığı (*GA*), tek bir evrenin ortalaması için veya evren ortalamaları arasındaki fark için hesaplanabilir. Öncelikle bir evren ortalaması için güven aralıklarının kullanılması üzerinde duracağız. Sonra iki evren ortalaması arasındaki farka ait güven aralıklarından bahsedeceğiz ve üç veya daha fazla ortalama söz konusu olduğunda aralıkların nasıl yorumlandığını tartışacağız.

Daha önce başka bir isim altında güven aralıklarına aşinalık kazanmış olabilirsiniz. Medyada, belirli bir katılımcı örneklemi ile yapılan tarama tipi araştırma sonuçlarını bildiren haberlere hiç rastladınız mı? Ve bazen bu haberlerde "hata payından" bahsedildiğini duydunuz mu? Kutu 11.2'de, hata payı kavramını ve bunun güven aralığı ile olan ilişkisini gözden geçirdik.

Tek Ortalama için Güven Aralıkları Evrene ait bir seçkisiz örneklemin ortalaması, evren ortalamasına ilişkin bir tahmin noktasıdır. Bununla birlikte, seçkisiz değişkenlik nedeniyle, örneklem ortalamaları arasında bir durumdan diğerine değişkenliğin olmasını bekleyebiliriz. Ortalamanın tahminî standart hatası ($s_{\overline{X}}$), örnekleme hatasının "normal" aralığına ilişkin bilgi verir. Güven aralığını hesaplarken, belirli bir güven derecesinde evren ortalamasını içine aldığını ifade ettiğimiz sayı aralıklarını belirleriz. Tahmin edebileceğiniz gibi, daha geniş bir aralık belirledikçe, ortalamanın aralığa dâhil olacağına dair güven düzeyimiz artmaktadır; fakat geniş aralıklar bize, evren ortalamasının kesin değerine ilişkin daha az spesifik bilgi verir. Araştırmacılar, %95'lik ve %99'luk güven aralıklarının, evren ortalamasının tahminine gereksinim duyulduğunda en iyi aralıklar olduğu üzerinde uzlaşmıştır.

Güven aralığı, evren ortalamasına ilişkin tahmin noktamızı ($\overline{X}$) merkez alacak şekildedir ve aşağıdaki formüller kullanılarak %95'lik güven aralığının sınırları hesaplanabilir:

$$\text{\%95'lik güven aralığının üst sınırı: } \overline{X} + [t_{.05}][s_{\overline{X}}]$$

$$\text{\%95'lik güven aralığının alt sınırı: } \overline{X} - [t_{.05}][s_{\overline{X}}]$$

KUTU 11.2

TARAMA TİPİ ARAŞTIRMA SONUÇLARINDA HATA PAYI

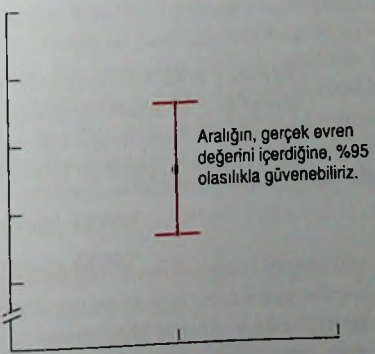
Bölüm 5'te öğrendiğiniz gibi, tarama tipi araştırmalar daha ziyade örneklemeye dayanır. Tarama tipi araştırmalar, bir evrenin özellikleri hakkında bilgi sahibi olmak istediğimiz zaman yapılmaktadır (örneğin; tercihler, tutumlar, demografik özellikler) ama çoğu zaman tüm evren üzerinde tarama yapmak olanaksızdır. Daha geniş bir evren hakkında bilgi edinmek için, örneklemden elde edilen sonuçlar kullanılır. İyi seçilmiş örneklemler, evrene ilişkin daha geçerli bilgiler verir; ama örnekleme ait bulguların, evreni kesin olarak tanımlaması pek olası değildir. Örneğin bir sınıfta 33 üniversite öğrencisinin yaş ortalaması 26.4 ise, sınıftan seçilen 10 öğrencinin yaş ortalamasının tam olarak 26.4 olması çok mümkün değildir. Benzer şekilde, bir şehrin nüfusunun %65'i var olan belediye başkanını destekliyorsa, %35'i de yeni bir belediye başkanı istiyorsa, şehir nüfusundan rastgele olarak seçilmiş 100 seçmenden oluşan bir örneklemde de mutlaka, aynı 65:35 oranını bekleyemeyiz. Sonuçlarda, örneklemeden dolayı bir miktar "kayma", asıl evren değerleri ile örneklemimizden elde ettiğimiz değerler arasında bir miktar "hata" bekleriz. Bu durumda önemli nokta, örneklemden elde edilen sonuçların, evreni ne ölçüde doğru temsil edebildiğidir.

Örneklem sonuçları ile gerçek evren değerleri arasındaki hata payını tahmin etmek mümkündür. Hata payı, evrene ilişkin bir değerin kesin bir tahminini vermek yerine (örneğin, "nüfusun %65'i var olan belediye başkanını desteklemektedir") gerçek evren değerini içerme olasılığı bulunan değerlerin aralığını verir (örneğin, "nüfusun %60ı ile %70'i arası, var olan belediye başkanını tercih etmektedir"). Peki bu aralık nedir?

Hata payı, örneklem sonuçları ile evrendeğerler arasında, tesadüfi faktörlere bağlı olarak ortaya çıkan farkın bir tahminidir. Hata payı bize, örnekleme hatası nedeniyle oluşabilecek değer aralıklarını verir –bir miktar hata bekleyebileceğimizi hatırlayın; evreni kesin olarak betimleyebilmeyi beklemeyiz. Çok sayıda seçmen ile bir kamuoyu yoklaması yapıldığını ve bir medya sözcüsünün şu sonuçları duyurduğunu varsayın: "Sonuçlar, örneklemin %63'ünün yönetimi desteklediğini göstermektedir ve %95'lik güven aralığı temel alındığında, kamuoyu yoklamasının %5'lik bir hata payı olduğunu söyleyebiliriz". Burada belirli bir güven aralığına göre (genellikle %95) rapor edilen hata payı, yönetimi destekleyen kişilerden oluşan asıl evrenin yüzdelerinin %58 ile %68 arasında tahmin edildiğini göstermektedir (örneklem değeri olan %63'e %5 eklenmekte ve bu değerden %5 çıkarılmaktadır). Bununla birlikte, gerçek evren değerini genellikle bilmediğimizi göz önünde bulundurmak gerekir. Örneklemden ve buna ilişkin hata payından hareketle elde ettiğimiz bilgi şu şekildedir: örneklemin %63'ü yönetimi desteklemektedir ve %95'lik bir güven aralığıyla söyleyebiliriz ki, eğer tüm evrene ulaşılabilseydi, evrenin %58'i ile %68'i arasındaki bir kısmı yönetimi destekleyecekti. Bu bulgu, örneklemden elde edilen değerin (%63) işaretlenmesi ve hata payını temsil eden hata çubuklarının eklenmesiyle, bir grafik üzerinde gösterilebilir. Şekil 11.1, örneklemden hareketle yapılan tahmin ve bunun etrafındaki hata çubuklarını göstermektedir.

Ulusal çapta yapılan tarama tipi araştırmalardaki hata payına ilişkin bilgiler medyada rutin olarak rapor edilmektedir. Bu araştırmaların amacı size "belirli bir hata payı ile birlikte", gerçek evren değerinin neye benzediğini söylemektir. Benzer biçimde, birçok bilimsel araştırmanın da amacı size, bir evren değerine ilişkin tahmine ait hata payını, şimdi daha sık kullanılan adıyla güven aralığını bildirmektir.

ŞEKİL 11.1 Hata çubukları, evrendeğerini tahmin etmedeki hata payını göstermek için kullanılmaktadır.



Örneklem ortalamasını ($\bar{X}$) ve ortalamanın tahminî standart hatasını ($s_{\bar{X}}$) hesaplamak için gerekli olan işlemleri açıklamıştık. Bu işlemlerde, %95'lik güven aralığının sınırlarına ait iki denklemde yer alan ve size tanıdık gelmeyen semboller *t* ve .05'tir.

Daha önce Bölüm 6'da .05'lik alfa (α) düzeyini kısaca tartışmıştık. Alfa düzeyi, istatistiksel anlamlılıkla ilgili çıkarımsal istatistikler ile (yani SHAT) ilişkilidir ve Bölüm 12'de alfa düzeyleri hakkında daha fazla bilgi vereceğiz. Güven aralıkları söz konusu olduğunda, oran olarak α = (1 – güven düzeyi) ifade edilmektedir. Başka bir deyişle, %95'lik güven aralığı için α = (1 – .95) = .05 ve %99'luk güven aralığı için α = (1 – .99) = .01'dir.

Denklemde bulunan *t* istatistiği ise serbestlik derecesini gösteren sayıya bağlı olarak tanımlanmaktadır ve *t*'nin istatistiksel anlamlılığı, ekte yer alan Tablo A.2'ye bakılarak belirlenebilir. Tek bir örneklem ortalaması için serbestlik derecesi $N - 1$'dir. Bölüm 12'de, SHAT'i tartıştığımız sırada, *t* istatistiği ile ilgili olarak daha fazla şey öğreneceksiniz. Şimdi yukarıda bahsedilen formülleri kullanarak güven aralığının hesaplanmasına ve uygun şekilde yorumlanmasına odaklanalım.

Bir örnekle, bir tek ortalama için güven aralığının nasıl elde edileceği gösterilebilir. Bir üniversitede seçkisiz yöntemle bir öğrenci örneklemi belirlediğinizi ve öğrencilerin zekâ düzeyini, kısa ama geçerli ve güvenilir bir yöntemle belirlediğinizi düşünün. Otuz öğrencinin (N = 30) test edildiğini ve ortalama zekâ puanının 115, örneklemin standart sapmasının da 14 olduğunu varsayın. Öğrencilere ait evren, üniversiteye devam eden binlerce öğrenciye karşılık gelecektir. Örneklem ortalaması da, evren ortalamasının iyi bir tahmin noktası (yeni evren ortalamasına ilişkin olarak yaptığımız en iyi tahmin) olduğu için, eğer farklı bir 30 kişilik seçkisiz öğrenci örneklemi belirlenseydi ve test edilseydi, örneklem ortalamasının tam olarak 115 olmayacağını bilmemiz gerekir. Bu seçkisiz süreç nedeniyle bir miktar kayma veya "hata" gerçekleşecektir. Standart hatanın, hata tahmini için bir ölçü olduğunu hatırlayınız.

Sadece evrene ait ortalamanın (μ) nokta tahminine ($\bar{X}$) dayanmak yerine, daha önce sunulan formülleri kullanarak evren ortalaması için %95'lik güven aralığını bulmakla bir aralık tahmini elde edebiliriz. Öncelikle ortalamanın tahminî standart hatası hesaplanır:

$$s_{\bar{X}} = \frac{s}{\sqrt{N}} = \frac{14}{\sqrt{30}} = \frac{14}{5.48} = 2.55$$

Sonrasında kritik *t* değerini elde ederiz. Örneklemde 30 öğrenci bulunduğu için, *t* istatistiğine ait serbestlik derecesi, 30 – 1 yani 29'dur. Tablo A.2'yi kullanarak, *t* değerinin .05 düzeyindeki alfa ve 29 serbestlik derecesi için 2.04 olduğunu görüyoruz. Güven aralığına ait formüller kullanılarak aşağıdaki sonuçlar elde edilir:

%95'lik güven aralığının üst sınırı = 115 + [2.04][2.55]

%95'lik güven aralığının alt sınırı = 115 − [2.04][2.55]

Üst sınır = 115 + 5.20 = 120.20

Alt sınır = 115 − 5.20 = 109.80

109.80 ile 120.20 aralığının %95'lik bir olasılıkla evren ortalamasını içerdiğini ("yakaladığını") söyleyebiliriz (bk. Kutu 11.3).

KUTU 11.3

TEK BİR ORTALAMA İÇİN GÜVEN ARALIKLARININ YORUMLANMASI: HALKALAR VE DİREKLER

Bir evren ortalaması için %95'lik güven aralığının hesaplanmasıyla ilgili olarak şunu söyleyebiliriz:

> elde edilen güven aralığının gerçek evren ortalamasını içerme olasılığı 95/100'dür.

Güven aralığı gerçek ortalamayı içeriyor veya içermiyor olabilir (örneğin; Mulaik, Raju ve Harsman, 1977). Ortalamaya ait güven aralığıyla ilişkilendiren .95'lik bir olasılık, eğer aynı boyuttaki farklı seçkisiz örneklemlere ilişkin bir çok güven aralığını bulmuş olsaydık, gerçek evren ortalamasının yakalanabilme olasılığına karşılık gelir. Başka bir deyişle, örneklem ortalamasını merkez alan güven aralıkları bize, eğer aynı çalışmayı aynı koşullar altında tekrarlasaydık sonucun ne olabileceğini söylemektedir (örneğin, Estes, 1997). Güven aralıklarımızı kullanarak, 100 tekrarlamanın 95'inde, gerçek ortalamayı yakalayabilmeyi bekleriz.

%95'lik güven aralığını hesaplamış olmak, şunu söyleyebilmek anlamına GELMEMELİDİR:

> gerçek ortalama bu aralığın içinde bir noktaya karşılık gelmektedir.

Bu ifade, başta kullanılan ifadeye benzer görünebilir fakat öyle değildir. Bulmaya çalıştığımız değerin değişken olmayan sabit bir değer olduğunu unutmayın; bu değer evrenin bir özelliği ya da parametresidir. Aralıklar ise sabit değildir; örneklenen verinin özellikleridir. Aralıklar, bir çalışmadan diğerine değişecek olan örneklem ortalamalarından ve yayılım ölçülerinden hareketle yapılandırılır ve dolayısıyla güven aralıkları da çalışmadan çalışmaya değişir.

Howell (2013), bu olguların güven aralıklarını yorumlama biçimimizle nasıl ilişkili olduğunu anlamaya yardımcı olmak için güzel bir benzerlik kurmuştur. Bir evrendeğeri (örneğin, evren ortalaması) bir direk biçiminde, güven aralığını ise direğe geçirmeye çalıştığımız halka biçiminde düşünmemizi önermiştir. Örneklem verisinden hareketle araştırmacı, direğe fırlatılan, belirli genişlikte halkalar yaratmaktadır. %95'lik güven aralığı kullanıldığında ise halka, denemelerin % 95'inde direğe geçecek, %5'inde direği kaçıracaktır. "Güven ile ilgili bir ifade, halkanın hedefi bulduğuna ilişkin bir olasılık ifadesidir; hedefin (evrendeğerin) halkaya isabet ettiğine ilişkin bir olasılık ifadesi değildir." (Howell, 2013; s. 194).

Aralık daraldıkça, aralığımız evren ortalamasını daha isabetli bir şekilde tahmin eder. Aralığın genişliğinin hem t istatistiğine hem de ortalamanın standart hatasına bağlı olduğunu, alt ve üst sınırlar için olan formülleri inceleyerek görebilirsiniz. Bu iki değer örneklem büyüklüğü ile ilişkilidir; şöyle ki, örneklem büyüklüğü arttığında her ikisi de azalır. Böyle olmakla birlikte, örneklem büyüklüğündeki artış en çok standart hatayı etkiler. Yukarıdaki örnekte örneklem büyüklüğünü iki katına çıkarmanın, 1.81 ($14/\sqrt{60}$) değerindeki bir standart hatayı, sonuç olarak, daha dar bir güven aralığını ortaya çıkaracağına dikkat edin. Kısaca, örneklem büyüklüğünü artırmak, ortalamaya ilişkin aralık tahminini iyileştirecektir.

İki Bağımsız Grup Ortalamasını Karşılaştırmak İçin Güven Aralıkları Ortalamalar arasındaki farka ilişkin güven aralıklarını oluşturmak için gerekli olan işlemler ile bunların arkasındaki mantık, tek ortalamadakiyle aynıdır. İlgilendiğimiz şey, şu an evren ortalamaları arasındaki fark (yani bağımsız değişkenin etkisi) olduğu için, $\overline{X}$ yerine $\overline{X}_1 - \overline{X}_2$ koyarız ve ortalamalar arasındaki farkın tahminî

standart hatasını kullanırız. İki evren ortalaması arasındaki fark için %95'lik güven aralığı aşağıdaki şekilde belirlenir:

$$GA(\%\ 95) = (\bar{X}_1 - \bar{X}_2) \pm (t_{0.5})(s_{\bar{X}_1 - \bar{X}_2})$$

Bu formüldeki t, Tablo A.2'de, $[(n_1 + n_2) - 2]$ değerine eşit olan serbestlik derecesi ve alfa = .05 için bulunabilir.

Ortalamalar arasındaki farkın tahminî standart hatası şu şekilde belirlenir:

$$s_{\bar{X}_1 - \bar{X}_2} = \sqrt{\left[\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}\right]\left[\frac{1}{n_1} + \frac{1}{n_2}\right]}$$

Örnek olarak, sözcük dağarcığı araştırmasına ait iki ortalama arasındaki fark için güven aralıklarını hesaplayalım. .05 düzeyindeki alfa ve 26 + 26 – 2 ya da 50 serbestlik derecesi için kritik t değeri Tablo A.2'den bulunmuştur. Bu değer, 2.009'dur. İki ortalama arasındaki farkın tahminî standart hatasını aşağıdaki şekilde elde edebiliriz:

$$s_{\bar{X}_1 - \bar{X}_2} = \sqrt{\left[\frac{(26 - 1)109.45 + (26 - 1)150.44}{26 + 26 - 2}\right]\left[\frac{1}{26} + \frac{1}{26}\right]} = 3.16$$

Buradan hareketle, evrene ait ortalamalar arası farkla ilgili % 95'lik güven aralığı için:

$$GA(\%95) = 18.46 \pm (2.009)(3.16)$$

$$= 18.46 \pm 6.35$$

Böylece üst sınır 18.46 + 6.35 = 24.81 ve alt sınır 18.46 – 6.35 = 12.11 olur. Bu nedenle, .95 düzeyinde bir güvenle 12.11 ve 24.81 aralığı, yaşlıları ile öğrencileri karşılaştırırken, sözcük dağarcığı testindeki doğru yüzdesi için gerçek evren farkını içermektedir. Dikkat ederseniz, sıfır değeri (0.0) bu aralık içinde yoktur. Bu, iki ortalama arasındaki fark için olan güven aralıklarını yorumlarken önemlidir (bk. Kutu 11.4). Eğer sıfır değeri aralık içerisinde yer alırsa, iki ortalama arasındaki gerçek farkın sıfır da olabileceği (yani farkın olmayabileceği) anlamına gelir (Cumming ve Finch, 2005). Bölüm 13'te güven aralıklarına dayanan bir analizi, araştırma raporunuzun bulgular bölümünde nasıl rapor edeceğinizi göstereceğiz.

Tekrarlı Ölçümler Deseninde İki Ortalamayı Karşılaştırmak İçin Güven Aralıkları Şimdiye kadar, iki bağımsız katılımcı grubunun dâhil olduğu deneyleri ele aldık. Farkında olduğunuz gibi deneyler ayrıca, her katılımcının deneyin her koşulunda yer almasını sağlayacak biçimde veya katılımcıların bağımlı değişken ile ilgili olan bazı ölçümler (örneğin, IQ puanı, ağırlık) açısından eşleştirilmesini sağlayacak biçimde de yapılabilir. Bu tür deneyler, eşleştirilmiş gruplar desenleri, grup içi desenler veya tekrarlı ölçümler desenleri olarak adlandırılır (bk. Bölüm 7). Örneğin bir bilişsel psikoloğun iki farklı bulmaca için bireylerin performansını karşılaştırmak istediğini düşünelim. İki farklı grupta yer alan bireylerden farklı bulmacalar üzerinde çalışmalarını istemek yerine araştırmacı, tek bir grup bireyden her iki bulmacayı da çözmeye çalışmalarını isteyebilir (deney materyalinin tekrarlı ölçümler deseninde sunumu için gerekli olan işlemler Bölüm 7'de açıklanmıştı). Sonrasında ise tüm katılımcılardan her iki bulmacaya

KUTU 11.4

İKİ ORTALAMA ARASINDAKİ FARK İÇİN GÜVEN ARALIKLARININ YORUMLANMASI: SIFIRI ARAMAK

İki ortalama arasındaki fark için %95'lik güven aralığını hesaplayınca şunu söyleyebiliriz:

> Elde edilen güven aralığının evrene ait gerçek ortalamalar arası farkı, yani mutlak etki büyüklüğünü içerme olasılığı 95/100'dür.

Güven aralığının genişliği, etki büyüklüğü hakkında bilgi verir. Güven aralıklarını kullanmakla, bağımsız değişkenimize ait olası etki büyüklüğüne ilişkin bilgi elde ederiz. Örneklemler ve işlemler değiştiği için, elde edilen etki büyüklükleri de çalışmadan çalışmaya değişir (örneğin bk. Grissom ve Kim, 2005). Güven aralığı "etki büyüklüğünün önem derecesine dair olası aralığı belirler" (Abelson, 1997, s. 130). Güven aralığı, etki büyüklüğünün, aralığın alt sınırı kadar küçük, üst sınırı kadar büyük olabileceğini gösterir. Araştırmacılar bazen bir etki büyüklüğünü yüksek derecede bir güven ile belirleyebilmek için ne kadar geniş bir aralığa ihtiyaç duyulduğunu gördüklerinde şaşırırlar (örneğin, Cohen, 1995). Bu nedenle güven aralığı daraldıkça, bağımsız değişkenimize ait gerçek etki büyüklüğünü tahmin etmede daha başarılı olmuş oluruz. Elbette güven aralığının genişliği, örneklem büyüklüğü ile doğrudan ilişkilidir. Örneklem büyüklüğünü artırarak, etkimizin tam olarak neye benzediği hususunda daha iyi bir fikir edinmiş oluruz.

Ortalamalar arası farka ait bir güven aralığının sıfırı içerip içermediğini belirlemek önemlidir. Eğer güven aralığı sıfırı içeriyorsa, iki evren ortalamasının farklılaşmadığına dair olasılığı kabul etmeliyiz. Bu durumda, bağımsız değişkenin bir etkisinin olduğu sonucuna varamayız. Güven aralıklarının etkimize ait olası bir aralığı verdiğini hatırlayın. Eğer sıfır olası değerler arasında ise bu durumda bir etkinin var olduğuna ilişkin belirsizliği kabul etmeliyiz. Bölüm 12'de bu durumun, SHAT'i kullanırken anlamlı olmayan sonuçların elde edildiği koşula benzediğini göreceksiniz.

ait puanlar elde edilecektir. İki puan arasındaki fark ise, tekrarlı ölçümler deseninde araştırmacının ilgilendiği ölçümdür.

Eşleşmiş gruplar veya tekrarlı ölçümler deseninde etki büyüklüğünü değerlendirmek için kullanılan işlemler, bağımsız gruplar deseni için gözden geçirdiklerimize kıyasla biraz daha karmaşıktır (bu durumlarda *d*'nin hesaplanmasına ilişkin bilgi için bk. Cohen, 1988; Rosenthal ve Rosnow, 1991). Getirilen önerilerden birisi, çalışmada bağımsız gruplar deseni kullanılmış gibi etki büyüklüğünü hesaplamak ve Cohen'in yönergesini (yani .20, .50, .80) uygulamaktır (örneğin, Zechmeister ve Posavac, 2003; ayrıca bk. Howell, 2013, ss. 203–204).

Güven aralıkları da iki koşulu olan tekrarlı ölçümler deseninde evren için ortalamalar arası farka ilişkin olarak hesaplanabilir. Bununla birlikte, altta yatan hesaplamalar böyle bir durumda değişecektir. Şöyle ki, her denek deneyin her iki koşulunda yer aldığında *t*, fark puanına dayanmaktadır (bk. Bölüm 12). Her denek için elde edilen iki puan birbirinden çıkarılarak, bir fark puanı elde edilmektedir. Fark puanlarının ortalaması ("D çizgi") şu şekilde belirlenir:

$$\overline{D} = \Sigma D/N$$

Burada D fark puanıdır, N ise fark puanlarının sayısıdır (yani puan çiftlerinin sayısı). $\overline{D} = \overline{X}_1 - \overline{X}_2$. olduğuna dikkat ediniz.

Fark puanlarının tahminî standart hatası ($S_{\bar{D}}$) şöyle belirlenir:

$$s_{\bar{D}} = \frac{s_D}{\sqrt{N}}$$ buradaki s_D, fark puanlarının standart hatasıdır.

t'nin kritik değerleri, ekte yer alan Tablo A.2'ye bakılarak, $N - 1$'e eşit olan serbestlik derecesi için elde edilebilir. Bu durumda N'nin, deneydeki katılımcıların veya puan çiftlerinin sayısına karşılık geldiğine dikkat ediniz.

Tekrarlı ölçümler desenindeki iki ortalama arasındaki fark için güven aralığı şöyle belirlenir:

$$GA = \bar{D} \pm (t_{05})(S_{\bar{D}})$$

İkiden Fazla Bağımsız Grup Ortalamasının Karşılaştırılması İçin Güven Aralıkları İkiden fazla ortalama söz konusu olduğunda sonuçların analizi ve yorumlanması amacıyla güven aralıklarının kullanımını göstermek için, bebeklerin "resimlerin niteliğini nasıl kavradıklarını" inceleyen bir çalışmayı ele alacağız (DeLoache, Pierroutsakos, Uttal, Rosengren ve Gottlieb, 1998). Bebeklerin, bir nesnenin resmi ile nesnenin kendisinin aynı şey olmadığını anlayıp anlamadıklarını hiç merak ettiniz mi? DeLoache ve arkadaşları, 5 aylık bebeklerin nesneler ile o nesnelere ait resimler arasındaki benzerliği tanıyabildiklerini, ama ikisinin aynı şey olmadığını da anladıklarını gösteren araştırmalar yaptı. Bununla birlikte bu araştırma bulguları bebeklerin ve küçük çocukların, resimlerdeki nesneleri yakalamaya ya da almaya çalıştıklarına, hatta resimdeki bir ayakkabıyı giymeye çalıştıklarına dair davranışlarıyla ilgili günlük bilgilerle pek uyuşmuyor. Bu tür günlük gözlemler bebeklerin ve çocukların, resimdeki iki boyutlu görüntüye rağmen nesnelerin resimlerine, gerçek nesneler gibi davrandıklarına işaret etmektedir. Dört çalışma ile DeLoache ve ark., "bebeklerin, resmi çizilmiş nesnelere ne ölçüde gerçek nesneler gibi davrandıklarını" (s. 205) araştırmıştır.

Burada, DeLoache ve ark. (1998) tarafından yürütülen dördüncü çalışmanın sonuçlarını ele alacağız. İlk üç çalışmada araştırmacılar şunları bulmuştur:

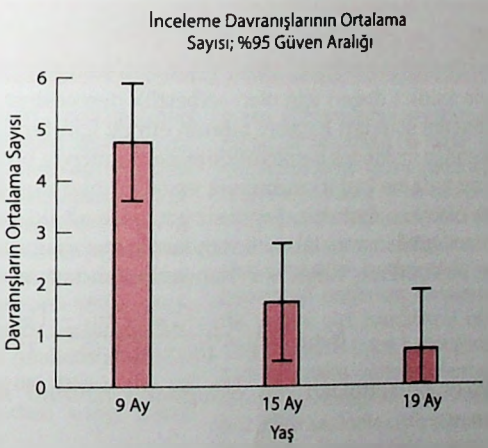
- dokuz aylık bebeklerin çoğu, "nesnelere ait (sık rastlanan plastik oyuncaklar) 8 adet çok gerçekçi renkli fotoğraf içeren resimli bir kitabı" incelerken, resmi çizilen bir nesneyi en az bir kez yakalamaya çalışmıştır (yakalama girişimlerinin sayısının ortalaması 3.7) (Çalışma 1);
- bebeklerin resimleri yakalamaya çalışmalarının nedeni, iki ve üç boyutlu nesneleri ayırt edememeleri değildir (Çalışma 2) ve
- "son derece yoksul ve çoğunlukla okur-yazar olmayan, bir Batı Afrika ülkesi olan Fildişi Sahili'nin kırsal bir bölgesinde yaşayan Beng ailelerin bebekleri" Amerikan bebeklerin yaptığı gibi resimleri (Beng topluluğunda yaygın olan nesnelerin resimlerini de içeren) elleri ile incelemişler ve yakalamaya çalışmışlardır (Çalışma 3) (bk. Şekil 11.2).

Dördüncü çalışmanın amacı ise çocukların resimlere olan davranışlarının yaşa bağlı olarak nasıl değiştiğinin belirlenmesidir.

ŞEKİL 11.2 Çocukların, resimlerin niteliğini nasıl kavradıkları meselesi, resmi çizilen nesneleri nasıl inceledikleri ve nasıl işaret ettikleri gözlenerek araştırılmıştır [Dr. Alma Gottlieb'nin Fildişi Sahili'nde, Beng çocukları ile yaptığı araştırma, DeLoache ve arkadaşlarından (1998) alınmıştır].



ŞEKİL 11.3 Dokuz aylık, 15 aylık ve 19 aylık çocukların inceleme davranışlarının, %95'lik güven aralıkları ile birlikte ortalama sayısı [Deloache ve arkadaşlarından (1998) izinle alınmıştır]



Çalışmada üç yaş grubu test edilmiştir: 9 ay, 15 ay ve 19 ay. Her grupta, 8'i kız, 8'i erkek olmak üzere 16 çocuk yer almıştır. Çocukların ellerini kullanarak resimleri inceleme davranışlarını (yakalama ve daha başka inceleme davranışları) gözlemenin yanı sıra, araştırmacılar resmi çizilen nesneleri işaret etme davranışlarını da kodlamışlardır. Çocukların inceleme davranışları ile ilgili sonuçlar Şekil 11.3'te verilmiştir.

Bağımsız değişken olan yaş, doğal gruplar desenidir ve üç düzeyi vardır: 9 ay, 15 ay ve 19 ay. Bu değişken, grafikte yatay eksende (x-ekseni) görülmektedir. Bağımlı değişken ise inceleme davranışlarıdır ve bu davranışların ortalama gözlenme sayısı dikey eksendedir (y-ekseni). Şekil 11.3'te görebileceğiniz gibi, inceleme davranışlarının ortalama sayısı en yüksek 9 aylık çocuklardadır; 15 ile 19 aylık çocuklarda ise çok daha düşük bir miktarda gözlenmiştir. Şekildeki bir başka önemli bilgi ise ortalamaların etrafındaki "çubuklar" ile ilgilidir. Bu çubukları, bağımsız değişken olan yaşın bir etkisinin olup olmadığına karar verirken kullanabiliriz.

Şekil 11.3'te her ortalamanın etrafındaki çubuklar, güven aralıklarını göstermektedir. Daha önce öğrendiğiniz gibi, güven aralıkları bize, bir evren değeri için beklediğimiz değerlerin aralığını göstermektedir. Örnekleme hatası nedeniyle evren değerini kesin olarak tahmin edemeyiz fakat olası değerlerin aralığını tahmin edebiliriz. Güven aralığımızda yer alan değerlerin aralığı daraldıkça, evren değerine ilişkin tahminimiz daha iyi olacaktır. Şekil 11.3'teki çubukların her birisi %95'lik güven aralığını göstermektedir. Bununla birlikte, çok sayıda grup için bu aralığın hesaplanması, sadece bir ortalama söz konusu olduğunda yapılan hesaplamaya göre biraz farklılık gösterir. Şöyle ki, ortalamanın tahminî standart hatasını hesaplarken çalışmadaki tüm gruplardan elde edilen birleştirilmiş varyansı kullanabiliriz. Bunun nasıl yapıldığı aşağıda gösterilmiştir.

%95'lik güven aralığı için olan formül, tek bir ortalama için olan ile aynıdır:

$$\%95\text{'lik güven aralığının üst sınırı: } \overline{X} + [t_{0.5}][s_{\overline{X}}]$$

$$\%95\text{'lik güven aralığının alt sınırı: } \overline{X} - [t_{0.5}][s_{\overline{X}}]$$

Öte yandan, $s_{\overline{X}}$'in hesaplanması, tek ortalama olan duruma göre farklıdır; dolayısıyla kritik t değeri için olan serbestlik derecesinin de hesaplanması değişir. Ortalamanın standart hatasını tahmin etmek için farklı gruplara ait varyansları, değişkenliğe ilişkin tek bir ölçü elde etmek amacıyla birleştirebiliriz. Bu durumda, çalışmamızda ne kadar grubumuz varsa, bunlara ait bilgiyi birleştiriyoruz. Yapılacak olan karşılaştırma, bağımsız gruplara ait iki veya daha fazla ortalamayı içeriyorsa, ortalamanın tahminî standart hatası aşağıdaki şekilde hesaplanır. Öncelikle, birleştirilmiş varyanstan hareketle standart sapma hesaplanır:[1]

$$s_{\text{birleşik}} = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2 + (n_3 - 1)s_3^2 + \cdots}{(n_1 - 1) + (n_2 - 1) + (n_3 - 1) + \cdots}}$$

Örneklem büyüklükleri eşit olduğunda tahminî standart hata şu şekilde hesaplanır:

$$s_{\overline{X}} = \frac{s_{\text{birleşik}}}{\sqrt{n}} \quad \text{burada, } n = \text{her bir grup için örneklem büyüklüğü}$$

Serbestlik derecesi ise $k(n - 1)$ şeklinde hesaplanır, burada k bağımsız grupların sayısıdır.

[1]Evrenin standart sapmasının birleşik tahmini, gruplar arası varyans analizindeki (ANOVA) ortalama hata karesinin kare köküne eşittir. Yani, $s_{\text{birleşik}} = \sqrt{\text{OKhata}}$. ANOVA'ya ilişkin tartışma için bakınız, Bölüm 12.

Tekrar Şekil 11.3'e baktığımızda 9 aylık bebekler için, inceleme davranışlarına ait örneklem ortalamasının 4.75 olduğu görülmektedir. Bu analizde, %95'lik güven aralığı için olan denklemdeki $[t_{05}]$ $[s_{\bar{X}}]$ ifadesi 1.14'e eşittir. Bu durumda, 3.61 ve 5.89 (4.75 ± 1.14) aralığının 9 aylık bebekler için evren ortalamasını içerdiğine %95 olasılıkla güvenebiliriz. Bu nedenle bu çalışmada, eğer daha büyük bir 9 aylık bebek örneklemi test edilecek olsaydı ortaya çıkacak olan inceleme davranışlarının ortalama sayısını tahmin etmek amacıyla, 9 aylık 16 bebekten oluşan örneklem kullanılmıştır. On beş aylıklar için inceleme davranışlarının ortalama sayısı 1.63'tür ve %95'lik bir olasılıkla, .49 ve 2.77 (1.63 ± 1.14) aralığının evren ortalamasını içereceğine güvenebiliriz. 19 aylıkların örneklem ortalaması .69'dur ve %95'lik güven aralığının alt sınırı 0.0 (izin verilen değerler tarafından sınırlanmıştır), üst sınırı ise 1.83'tür (.69 ± 1.14).

Kutu 11.5'te, üç ve daha fazla ortalama söz konusu olduğu zaman güven aralıklarının nasıl yorumlanacağı hakkında bilgi verilmektedir.

Araştırma sonuçlarına ait çizilen grafiklerdeki "çubukların" incelenmesi ile ilgili olarak son bir uyarıda bulunmak gerekmektedir. Dergi makalelerinde veriye ilişkin grafiklerde gösterilen çubuklar bazen güven aralıklarını göstermekte fakat ayrıca ortalamanın standart hatasını veya standart sapmayı da gösterebilmektedir (Cumming ve Finch, 2005). (%95'lik güven aralığının

KUTU 11.5

ÜÇ VEYA DAHA FAZLA ORTALAMA SÖZ KONUSU OLDUĞUNDA GÜVEN ARALIKLARININ YORUMLANMASI: ARALIKLAR ÖRTÜŞÜR MÜ?

Birçok araştırmada evren ortalamasının kendi değerini tahmin etmekle ilgilenmeyiz. Örneğin 9 aylık çocukların resimleri ortalama kaç kez yakalamaya çalışacaklarını beklememiz, bizi çok ilgilendirmez. Bunun yerine bizi ilgilendiren şey, evren ortalamalarının örüntüleri ve evren ortalamaları arasındaki ilişkilerdir (Loftus ve Masson, 1994). Yani, farklı grupların davranışlarını karşılaştırabiliyor olmayı umarız. Bu da güven aralıklarının karşılaştırılmasıyla yapılabilir. DeLoache ve arkadaşlarının çalışmasındaki veriyi tekrar ele alalım.

Şu soruyu sormak için evren ortalamalarına ilişkin tahminlerimizi kullanabiliriz: Farklı yaş gruplarından bebekler, farklı sayıda inceleme davranışı göstermekte midir? Bu soruyu yanıtlamak için Şekil 11.3'teki güven aralıklarında gözlenen örtüşmeleri inceleyebiliriz. Güven aralığının olasılıkla ilgili bir şey olduğunu (örneğin, .95) hatırlayınız. Aralık, evren ortalamasını içermektedir; aralığın genişliği, tahminimizin ne kadar kesin olduğunu göstermektedir. Güven aralıklarının bir evren değerini ne kadar iyi bir şekilde tahmin ettiğimiz hakkında bilgi verdiğini unutmayın. Güven aralıkları, farkların "istatistiksel olarak anlamlı" olup olmadığını inceleyen, iki veya daha fazla ortalamayı doğrudan karşılaştıran, *t*-testi ya da *F*-testi gibi istatistiksel bir test değildir. Yine de, daha önce belirttiğimiz gibi, araştırmacılar sıklıkla evren ortalamalarının nasıl bir örüntü gösterdiği ile ilgilenmektedirler ve bu örüntüleri saptamak için güven aralıklarını kullanabiliriz. Aralıklar örtüşmediği zaman evren ortalamalarının farklılaştığına karar verebiliriz. Örtüşmeyen aralıklar, örneklem ortalamaları tarafından tahmin edilen evren ortalamalarının muhtemelen aynı olmadığını söylemektedir. Örneğin, 9 aylık bebekler için olan %95'lik güven aralığı, 15 aylık bebeklerin aralığı ile örtüşmemektedir. Buradan hareketle, 9 aylık bebeklerin, resimlere bakarken yaptıkları

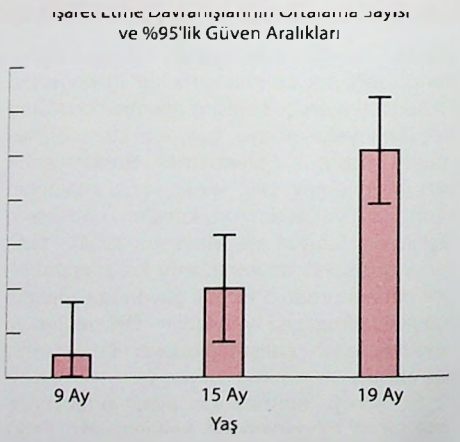
(devam ediyor)

inceleme davranışlarının miktarı açısından 15 aylık bebeklere göre farklılaştığı sonucuna varabiliriz (örneklem ortalamaları 9 aylık bebeklerin resimleri, 15 aylık bebeklerden daha fazla incelediğini göstermektedir). Benzer bir sonuca, 9 aylık ve 19 aylık bebeklere ait aralıklar karşılaştırıldığı zaman da varılabilir. On beş ve 19 aylık bebeklere ait aralıkların karşılaştırılmasından ise daha farklı bir sonuç çıkarılmalıdır. Şekil 11.3, bu iki grubun aralıklarının örtüştüğünü göstermektedir. Bundan nasıl bir sonuç çıkarmalıyız? Eğer aralıklar az örtüşüyorsa, ortalamalar arası gerçek farkla ilgili belirsizliği kabul etmeli ve farklılığa dair bir yargıya varmamalıyız. Eğer aralıklar, gruplardan birinin örneklem ortalamasının, diğer grubun aralığının içinde yer alacağı şekilde örtüşüyorsa, evren ortalamalarının farklılaşmadığı sonucuna varabiliriz (bk. Zechmeister ve Posavac, 2003). Cumming ve Finch (2005), örtüşen aralıkların yorumlanmasıyla ilgili olarak, oranlı örtüşmeyi temel alan daha ayrıntılı bir analiz önermiştir.

Bu yönergelerden hareketle, Şekil 11.3'te gösterilen 15 ve 19 aylık bebekler arasındaki fark için nasıl bir sonuç çıkarabiliriz? Şekilde de görüldüğü gibi, %95'lik aralıklar, 15 aylık bebeklerin güven aralığı 19 aylık bebeklerin örneklem ortalamasını içerecek şekilde örtüşmektedir. Bu nedenle evren ortalamalarının farklılaşmadığını söyleyebiliriz. Örneklem ortalamaları farklı olsa bile (sırasıyla 1.63 ve .69), evren ortalamalarının farklılaştığı sonucunu çıkaramayız (ayrıca psikolojideki araştırmalarda evreni tanımlamakla, örneklemi tanımlamaktan daha çok ilgileniriz). Örneğin, Şekil 11.3'teki aralıkların çakışması göz önüne alındığında, 15 aylık çocukların gerçek evren ortalamasının .69 (19 aylık bebeklerin örneklem ortalaması olan değer) olması ihtimal dâhilindedir. Şekil 11.3'te sunulan ortalamalardan ve güven aralıklarından hareketle, 9 aylık bebeklerin resimleri elleri ile, 15 ve 19 aylık bebeklere göre, daha fazla incelediği, ayrıca yaşı daha büyük olan bu iki grubun gösterdikleri inceleme davranışlarının sayısı açısından farklılaşmadığı sonucunu çıkarabiliriz. Yaşı daha büyük olan iki grubun (evrenlerin) farklılaşma olasılığının olmadığını söylemediğimize dikkat edin. Buradaki veriler göz önünde bulundurulduğu zaman, bir farkın var olduğunu söyleyemeyiz; bununla birlikte veri, farkın var olmadığını da kesin olarak söylememektedir. Evren ortalamalarının daha hassas bir tahminini yapmak üzere belki daha geniş örneklemlerin kullanıldığı daha fazla araştırma yapılana kadar beklemeliyiz.

Bebeklerin resimleri işaret etmesi ile ilgili olarak DeLoache ve arkadaşlarının verisi Şekil 11.4'te sunulmuştur. Bu şekildeki ortalamaları ve güven aralıklarını temel alarak nasıl bir sonuca varırsınız? On beş aylık bebekler, işaret etme davranışı açısından 9 aylık bebeklerden farklılaşmakta mıdır? Neden farklılaşmakta veya farklılaşmamaktadır? On dokuz aylık bebekler, 15 aylık bebeklerden farklılaşmakta mıdır, neden farklılaşmakta veya farklılaşmamaktadır?

ŞEKİL 11.4 Dokuz aylık, 15 aylık ve 19 aylık çocukların, işaret etme davranışlarının, %95'lik güven aralığı ile birlikte ortalama sayısı



yaklaşık olarak hesaplanması ile ilgili hızlı bir teknik, ortalamanın standart hatasının 2 ile çarpılmasıdır.). Yazarlar bazen, işleri daha da karmaşıklaştırarak, neyin sunulduğu konusunda okuyucuları bilgilendirmemektedir. Eğer çubuklar sunuluyorsa, bunların neyi ifade ettiği ve nasıl hesaplandığı konusunda okuyucuları bilgilendirmek önemlidir (Estes, 1997).

ÖRNEK: KORELASYONEL ÇALIŞMA İÇİN VERİ ANALİZİ

- Aynı kişi, olay ya da başka bir şeye ait iki faklı ölçüm birlikte değişiyorsa korelasyonun varlığı söz konusudur –başka bir deyişle bir değişkendeki puanlar, diğer değişkendeki puanlar ile birlikte değişmektedir.

Bölüm 2'de gördüğünüz gibi kestirim (prediction) bilimsel yöntemin önemli amaçlarından birisidir. Korelasyonel araştırmalar genellikle kestirim için temel teşkil eder. Aynı kişi, olay ya da başka bir şeye ait iki faklı ölçüm birlikte değişiyorsa korelasyonun varlığı söz konusudur –başka bir deyişle bir değişkendeki puanlar, diğer değişkendeki puanlar ile birlikte değişmektedir. Örneğin, sigara içme ve akciğer hastalığı arasında iyi bilinen bir ilişki mevcuttur. Bireyler daha fazla sigara içtikçe (örneğin, sigara içme süresi ile ölçülebilir), akciğer hastalığına yakalanma olasılığı artar. Bu nedenle sigara içme ve akciğer hastalığı birlikte değişmekte veya birlikte hareket etmektedir. Bu korelasyon aynı zamanda şu şekilde de ifade edilebilir: Bireyler daha az sigara içtikçe, akciğer hastalığına yakalanma olasılıkları düşer. Bu korelasyonu temel alarak akciğer hastalığına ilişkin kestirimde bulunabiliriz. Örneğin, bir bireyin ne kadar süredir sigara içtiğini biliyorsak, bu kişinin akciğer hastalığına yakalanma olasılığını kestirebiliriz (bir dereceye kadar). Tahminlerimizin doğası ve bu tahminleri yaparkenki güven düzeyimiz, korelasyonun yönüne ve gücüne bağlıdır.

Korelasyonel analizler sıklıkla tarama tipi araştırmalar ile ilişkilendirilir (bk. Bölüm 5). Bu araştırmalarda katılımcılar, demografik değişkenlerin (örneğin, yaş, gelir) yanında kendilerinin tutumları, görüşleri ve psikolojik iyi oluşları ile ilgili sorulan anket sorularını cevaplar. Sonrasında ise araştırmacı farklı yanıtların arasında bir ilişki olup olmadığını, yani aralarında korelasyon olup olmadığını inceler. Düşük benlik saygısı olduğunu bildiren bireyler, aynı zamanda

İLERİ ALIŞTIRMA
GÜVEN ARALIKLARI İLE İLGİLİ BİLGİNİZİN SINANMASI

Güven aralıklarının, genellikle istatistiksel anlamlılık testlerinin yorumlaması ile ilgili, özellikle sıfır hipotezi anlamlılık testi (SHAT) ile ilgili bazı ortak sorunları vardır. Yine de veri analizine güven aralıkları dâhil edilebilir ve edilmelidir de. Güven aralıklarını doğru bir biçimde kullandığınızdan emin olmak için, bu analiz tekniği ile ilgili bilginizi ölçen aşağıdaki testi hazırladık.

Üç düzeyi (A, B, C) olan bir bağımsız değişkenin davranış üzerindeki etkisini incelemek için bağımsız gruplar deseninin kullanıldığını varsayın. Burada, her koşula atanmış 15 katılımcı vardır ve her koşul için merkezî eğilim ölçüleri ile değişkenlik ölçüleri belirlenmiştir. Ayrıca araştırmacı, her ortalama için %95'lik güven aralığını yapılandırmıştır. Buradan hareketle araştırmacı aşağıdaki sonuçlara varırsa, bunlardan hangileri doğru, hangileri yanlış olur?

1 Güven aralığının genişliği evren ortalamalarının ne kadar kesin olarak tahmin edildiğini gösterir.
2 İki aralık örtüşüyorsa, evren ortalamalarının aynı olduğundan emin olabiliriz.
3 Gerçek evren ortalamasının her bir aralığa denk gelme olasılığı %95'tir.
4 İki aralık örtüşmüyorsa evren ortalamalarının farklı olduğuna dair olasılık %95'tir.
5 İki aralık örtüşmüyorsa evren ortalamalarının farklı olduğuna dair güçlü bir kanıtımız vardır.

flört etmekte zorlandıklarını da bildirmekte midirler? Çocukların kreşte geçirdikleri zamanın uzunluğu, annelerine olan bağlanmalarına dair ölçümler ile ilişkili midir? SAT puanları, üniversiteden sonraki başarıyı tahmin etmekte midir?

Araştırmacıların korelasyonel çalışmaları nasıl analiz ettiği ve yorumladığı aşağıda ele alınmıştır.

Aşama 1: Verileri Tanımaya Başlama

Korelasyonel bir çalışmada her zaman iki küme puan yer aldığı ve ilgilenilen esas şey bu puanlar arasındaki ilişki olduğu için, veri analizinin aşamaları, çalışmanın merkezinde ortalamaların karşılaştırılmasının olduğu durumdan biraz daha farklı ilerler. Örnek olarak, varsayın ki bir araştırmacı üniversite öğrencilerinden öz bildirim (self report) ile alınan psikolojik iyi oluş ile ilgili iki ölçümü ilişkilendirmek istiyor (öz bildirim ile ilgili veriye ait tartışma için bk. Bölüm 5). Her iki ölçüm de 10'lu puanlama ölçeği biçimindedir. Ölçümlerden biri "Notlar ile ilgili ne kadar endişeleniyorsunuz?" (1 = *hiç*, 10 = *çok fazla*) sorusuna dayanmaktadır. İkinci ölçüm ise "Sınavlar sırasında konsantre olmada ne kadar zorluk çekiyorsunuz?" (1 = *hiç*, 10 = *çok fazla*) sorusuna dayanmaktadır.

Verilerin Temizlenmesi Her katılımcıya ait iki puan vardır ve bu iki puan kümesi, mümkün olmayan puanlar (örneğin, ölçeğin aralığı dışındaki sayılar) gibi hatalar için olduğu kadar, aşırı değerler için de dikkatli bir biçimde kontrol edilmelidir. Her kümedeki verinin incelenmesi için gövde ve yaprak gösterimi kullanılabilir. Ölçek kullanıldığında tipik şekilde yapıldığı gibi, olası tepkiler sınırlandırıldığı zaman aşırı değerlerin gözlenme olasılığı, yanıtlara sınır konmadığı duruma göre (örneğin, yıllık gelirin rapor edilmesi) daha azdır.

Sonuç Araştırmacı, ancak verinin hata veya bulguları çarpıtabilecek değerler içermediğinden emin olduktan sonra analize devam etmelidir.

Aşama 2: Verileri Özetleme

- Korelasyonel veriler için temel betimsel yöntemler, dağılım grafiğinin hazırlanması ve korelasyon katsayısının hesaplanmasıdır.
- Korelasyonun büyüklüğü veya derecesi, dağılım grafiğinde noktaların düz bir çizgi üzerine ne kadar iyi denk geldiği belirlenerek görülebilir; daha güçlü korelasyonlar daha açık bir şekilde, noktalardan oluşan düz bir çizgiye benzer (doğrusal eğilim).
- Bir korelasyon katsayısının büyüklüğü, –1.0 (mükemmel negatif ilişki) ile +1.0 (mükemmel pozitif ilişki) arasında değişir; korelasyon katsayısının 0 olması, bir ilişkinin olmadığı anlamına gelir.

Verilerin özetlenmesi, her puan kümesi için betimsel istatistiklerin incelenmesiyle başlar. Daha sonra bu puan kümeleri arasındaki ilişkinin derecesi hem grafiksel hem de sayısal olarak özetlenir.

Merkezî eğilim ve Değişkenlik Merkezî eğilim ve değişkenlik ölçüleri iki puan kümesi için de hesaplanmalıdır. Hipotetik çalışmamızdaki her iki yanıt kümesi için ortalamalar ve standart sapmalar şöyledir:

	Endişe	Konsantrasyon Güçlüğü
ort.	5.45	5.30
S	1.93	1.98

Temel Kavram

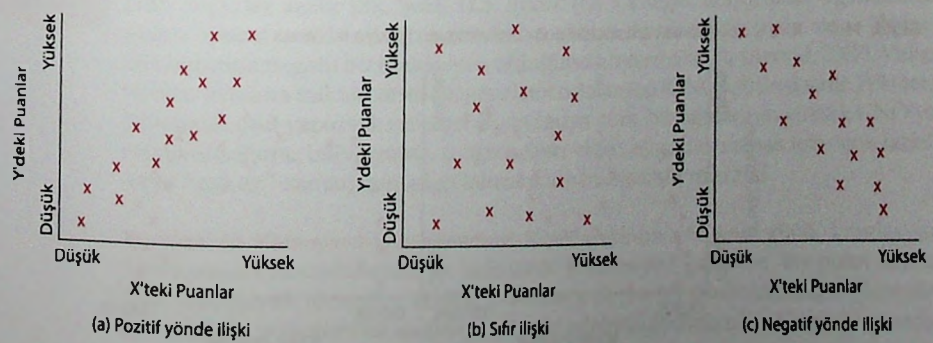
Korelasyonel bir çalışmada ilgilendiğimiz esas şey ortalamalar arasındaki fark değil, puan kümeleri arasındaki ilişkidir. Korelasyonel veriler için ana betimsel teknikler dağılım grafiğinin hazırlanması ve korelasyon katsayısının hesaplanmasıdır. **Dağılım grafiği**, iki puan kümesi arasındaki ilişkiyi gösterir.

Dağılım Grafiğinin Çizilmesi Her bireyin, her değişken için (örneğin, endişe puanları ve odaklanma güçlüğü) bir değeri (veya puanı) vardır. İki değişken için olan puanlar x-ekseninde ve y-ekseninde gösterilmektedir. Dağılım grafiği, her puan çifti için kesişme noktasını göstermektedir. Korelasyonun büyüklüğü veya derecesi, dağılım grafiğinde noktaların düz bir çizgi üzerine ne kadar iyi denk geldiği belirlenerek görülebilir; daha güçlü korelasyonlar daha açık bir biçimde noktalardan oluşan düz bir çizgiye benzer. Şekil 11.5, üç farklı dağılım grafiğini göstermektedir. Birinci ve üçüncü grafikte korelasyon, ikincisine göre daha güçlüdür çünkü noktalar düz bir çizgiye daha yakındır.

Yirmi üniversite öğrencisinin yukarıda tanımladığımız sorulara yanıt verdiğini varsayın. Ayrıca verinin hatalar ile diğer başka anormallikler açısından dikkatlice incelendiğini ve verinin temiz olduğuna karar verildiğini de varsayın.

Burada bir ölçümdeki puanların, ikinci ölçümdeki puanlarla ilişkili olup olmadığını (yani birlikte hareket edip etmediğini) bulmak istiyoruz. Notlar ile ilgili olarak katılımcıların bildirdiği endişe, sınavlarda konsantre olma güçlüğü ve öz bildirim raporları ile ilişkili midir? Dağılım grafiği hazırlanırken,

ŞEKİL 11.5 İki değişkene ait puanlar arasındaki pozitif (a), sıfır (b) ve negatif (c) korelasyonu gösteren üç dağılım grafiği



geleneksel olarak, "ilk gelen" davranış ölçümü veya ikinci davranışı kestirmek için kullanılan davranış, yatay *x*-eksenine yerleştirilir. İlki tarafından kestirilen ikinci davranış ise yatay *y*-eksenine yerleştirilir. Çoğu durumda bu kararı vermek kolaydır. Eğer gönüllü katılımcıların kanlarındaki alkol düzeyi ile sürüş simülatöründeki performanslarını ilişkilendiriyor olsaydınız, önce alkolün tüketileceğini, daha sonra simülatördeki sürüş performansının ölçüleceğini kolayca görebilirdik. Burada, kandaki alkol düzeyi, sürüş simülatöründeki performansı kestirmek için kullanılacaktır. Diğer bazı durumlarda ise böyle bir kararı vermek bu kadar kolay olmayabilir. Örneğin, ders notlarına ilişkin endişe, sınavlardaki odaklanma güçlüğünden önce mi gelmektedir? Ya da sınavlarda yaşanan odaklanma güçlüğü, notlara ilişkin endişeyi mi doğurmaktadır? Böyle bir durumda her ikisinin de geçerli olabileceğini kabul ederiz.

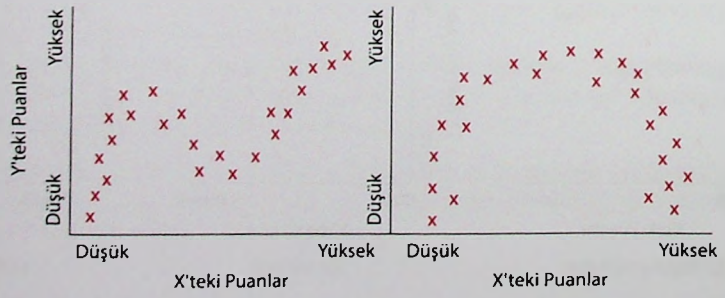
Temel Kavram

Sonraki aşamada, dağılım grafiğini olası eğilimler için inceleriz. Daha açık söylemek gerekirse, dağılım grafiğinde **doğrusal bir eğilimin** varlığına dair kanıt olup olmadığını araştırırız. Basitçe doğrusal eğilim, düz bir çizgi ile özetlenebilecek bir eğilim türüdür. Gördüğünüz gibi Şekil 11.5'teki (a) ve (c) dağılım grafikleri doğrusal bir eğilime kanıt teşkil etmektedir. Dağılım grafiğinde herhangi bir eğilim görememek de mümkündür. Bu durumda ölçümlerden birindeki puanlara, ikinci ölçümdeki düşük, orta ya da yüksek düzeydeki puanların herhangi birisi eşlik etmektedir. Şekil 11.5'in orta kısmında olduğu gibi, grafikte fark edilebilir bir eğilim yoksa, puan kümeleri arasında bir ilişki olmadığı sonucuna varabiliriz. Bu durumda ölçümlerden birindeki puanlar ile ilgili bilgimizi, ikinci ölçümdeki puanları tahmin etmek için kullanamayacağımıza dikkat edin.

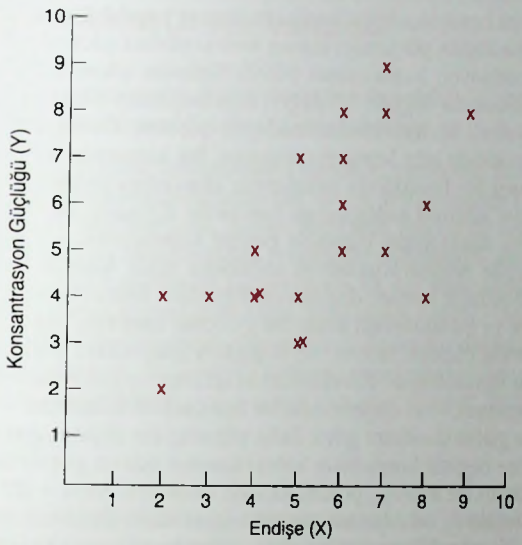
Son olarak, dağılım grafiğinde doğrusal olmayan bir ilişki görmek de mümkündür. Şekil 11.6'da, değişkenler arasında doğrusal olmayan iki ilişki örneği gösterilmektedir. Bu ilişkilerin ilginç olduğuna ve hatta daha fazla inceleme yapmaya değer olduğuna karar verebiliriz; bununla birlikte, doğrusal olmayan bir ilişki, korelasyon katsayısının yorumlanması ile ilgili olarak ciddi problemleri beraberinde getirir. Sonuçta, dağılım grafiğindeki ilişki doğrusal olmayan bir ilişki ise korelasyon katsayısı hesaplanmamalıdır. Aynı zamanda, dağılım grafiğindeki uç değerler de korelasyon katsayısının yorumlanmasında sorunlara neden olur.

Şekil 11.7, hipotetik araştırmamızdaki ölçümler olan endişe (X) ve odaklanma güçlüğü (Y) puanları arasındaki ilişkiyi göstermektedir. Buradaki durumda hangi faktörün "önce geldiğini" bilemediğimiz için, Şekil 11.7'deki dağılım grafiğinde endişe ölçümünü *x*-eksenine ve odaklanma güçlüğü ölçümünü

ŞEKİL 11.6 X ve Y değişkenleri arasındaki doğrusal olmayan ilişkiye iki örnek



ŞEKİL 11.7 Notlar hakkındaki endişe derecesi ile ilgili öz bildirim raporu ölçümü puanları (X) ile sınavlar sırasında yaşanan odaklanma güçlüğüne ilişkin öz bildirim raporu ölçümü (Y) arasındaki ilişkiyi gösteren dağılım grafiği. Grafikteki her bir nokta, her katılımcıya ait iki ölçümün kesişimini göstermektedir



y-eksenine rastgele koyduk. Yani, endişe ölçümünü, odaklanma ölçümünü tahmin etmek için kullanıyoruz.

Temel Kavram

Korelasyonun yönü pozitif veya negatif olabilir. **Pozitif korelasyon,** ölçümlerden birindeki değerler arttıkça, diğer ölçümdeki değerlerin de artacağını göstermektedir. Şekil 11.7'de gösterilen bu türden bir ilişkidir. Endişe puanları arttıkça, odaklanma güçlüğü puanları da artmaktadır. Öğrencilerin, notlar hakkında daha fazla endişelendikçe, sınav sırasında daha fazla odaklanma güçlüğü yaşayacaklarını kestirebiliriz. Korelasyonlardan hareketle ters yönde bir kestirim de yapılabilir. Bir kimsenin sınavlarda odaklanma güçlüğü yaşadığını biliyorsak, o kimsenin notlar hakkında endişe duyma olasılığı olduğunu da "tahmin" edebiliriz.

Temel Kavram

Negatif korelasyonda, ölçümlerden birindeki değerler arttıkça, diğer ölçümdeki değerler azalır [bk. Şekil 11.5, grafik (c)]. Örneğin, araştırmalar öğrencilerin video oyunu oynamak için harcadıkları zaman ile akademik performans ölçümleri arasında negatif bir korelasyon olduğunu göstermektedir (Anand, 2007). Video oyunu oynama miktarı arttıkça, genel not ortalaması (GNO) düşmektedir. Peki ters yönde tahmin yaparsak ne olur? Bu bulguya göre, bir öğrencinin yüksek GNO'ya sahip olduğunu biliyorsanız, o öğrencinin video oyunu oynamak için "çok fazla" veya "çok az" zaman harcadığı tahmininde bulunabilir misiniz?

Korelasyon Katsayının Hesaplanması Korelasyonun yönü ve gücü, korelasyon katsayısının hesaplanması ile belirlenir. Korelasyon katsayısı, bir puan kümesini kullanarak (örneğin, endişe puanları), başka bir puan kümesini (örneğin, odaklanma puanları) ne kadar iyi tahmin edebildiğimizin nicel bir göstergesidir.

Korelasyon katsayısı, iki değişken arasındaki ilişkiyi, hem ilişkinin yönü, hem de ilişkinin büyüklüğü açısından ifade eder. En sık kullanılan korelasyon katsayısı, Pearson Momentler Çarpımı Korelasyon Katsayısıdır ve *r* ile gösterilmektedir. Elektronik bir hesap makinesi ile veya bir bilgisayar programı ile kolayca hesaplanabilir (İnternette arama yapıldığında, korelasyonun hesaplanmasına ilişkin yöntemler sunan web sayfaları çıkacaktır.).

Korelasyon katsayısının büyüklüğünün (derecesinin) mutlak değeri 0 ile 1.00 arasında değişir. 0.0 değeri korelasyonun olmadığını ve kestirim yapmak için yeterli bir temel bulunmadığını gösterir. Örneğin, zekâ ve zihinsel bozukluk arasında sıfır korelasyon vardır; bir kimsenin IQ'sunu bilmekle, o kişinin zihinsel bir bozukluğa yakalanma olasılığını tahmin edemeyiz (aynı şekilde, kişinin zihinsel sağlığından hareketle IQ'sunu da tahmin edemeyiz). +1.00 değeri, mükemmel düzeyde pozitif korelasyon, –1.00 değeri ise mükemmel düzeyde negatif korelasyon anlamına gelir. Korelasyon katsayısı +1.00 veya –1.00 olduğu zaman, dağılım grafiğindeki bütün noktalar düz bir çizgi üzerine düşer ve bu durumda kesin bir güvenle kestirim yapabiliriz. Katsayı 0 ile 1.00 arasında olduğu zaman, iki değişken arasındaki ilişkinin orta düzeyde olduğunu kestirebiliriz. Korelasyonun işaretinin, onun sadece yönünü gösterdiğini unutmayın; – .46 değerindeki bir korelasyon katsayısı, + .20 değerindekine göre daha güçlü (kestirim gücü daha yüksek) bir ilişkiyi gösterir (*Not:* Uygulamada sadece negatif korelasyon katsayılarının işareti gösterilir; artı veya eksi işareti olmayan bir katsayı, pozitif katsayı demektir, yani + .20 = .20).

Endişe ve odaklanma güçlüğü arasındaki ilişkinin araştırıldığı hipotetik çalışmamızda, 20 öğrenciden elde edilen korelasyon katsayısı .62'dir. Korelasyon katsayısı, dağılım grafiğinde gözlediğimiz şeyin nicel bir göstergesini sunmaktadır. Buradaki iki değişken arasında pozitif bir korelasyon olduğunu söyleyebiliriz: Öğrencilerin endişe düzeyi arttıkça, sınav sırasında odaklanma güçlüğü yaşama olasılıkları da artmaktadır. Fakat endişenin, öğrencilerin odaklanma güçlüğü yaşamalarına neden olduğunu söyleyebilir miyiz?

Korelasyon ve Nedensellik Bölüm 4 ve 5'te korelasyonlara ilişkin yaptığımız tartışmadan hatırlayabileceğiniz gibi, "korelasyon, nedenselliğe işaret etmez". İki değişken arasında korelasyon bulunduğunu biliyor olmak, bu değişkenlerden birisinin diğerine neden olduğu çıkarımını yapmamıza izin vermez (zamansal olarak biri diğerini takip ediyor olsa bile). Düşük not alma endişesi sınava odaklanma güçlüğüne neden olabilir veya sınava odaklanma güçlüğü, düşük not alma endişesine neden oluyor da olabilir. Dahası, düşük not alma endişesi ile sınava odaklanma güçlüğü arasındaki ilişkinin nedeni üçüncü bir değişken de olabilir; buna sahte ilişki denir. Örneğin, 'ders çalışma süresi' değişkeni, düşük not alma endişesi ile sınava odaklanma endişesi arasında gözlenen ilişkiyi açıklayan üçüncü bir değişken olabilir. Ders çalışma süresi azaldıkça, öğrenciler daha fazla not endişesi yaşıyor ve sınava odaklanmada daha fazla zorluk çekiyor olabilir.

Sonuç Pearson Momentler Çarpımı Korelasyon Katsayısı, iki değişken arasındaki ilişkiyi özetlemek için kullanılabilir. Bununla birlikte, Pearson *r*'nin hesaplanmasından önce iki değişkene ait dağılım grafiği incelenmelidir, çünkü değişkenler arasındaki ilişkinin en uygun şekilde düz bir doğru ile özetlenip özetlenemeyeceğinden, yani doğrusal bir eğilimin olup olmadığından emin olmak gerekir. Korelasyon

katsayısı 1.00'a (ya da –1.00'a) yaklaştıkça, dağılım grafiğinde gözlenen iki değişken arasındaki ilişki düz bir çizgiye yaklaşır ve değişkenlerden birisini, diğerini temel alarak kestirme becerimiz artar.

Aşama 3: Korelasyon İçin Güven Aralığının Hesaplanması

- Evren ortalaması olan μ için yaptığımıza benzer şekilde, evren korelasyonu olan ρ'ye ait bir güven aralığı tahmini elde edebiliriz.

Örneklemden hareketle hesaplanan Pearson r, evrendeki korelasyonun bir tahminidir; tıpkı örneklem ortalamasının, evren ortalamasının (μ) bir tahmini olması gibi. Evren korelasyonu, Yunan harfi olan rho (ρ) ile sembolize edilir. Örnekleme hatası, yani örneklemden örnekleme değişkenlik, örneklem ortalamasında olduğu gibi korelasyon katsayısında da söz konusudur. Bu nedenle bazı durumlarda evren değeri olan ρ için, tıpkı evren değeri olan μ için olduğu gibi, bir aralık tahmini elde etmek isteyebiliriz. Başka bir deyişle, ρ için güven aralığını hesaplayabiliriz. Bu konuya değinmeyeceğiz fakat daha ayrıntılı istatistiksel işlemler sunan kitaplar mevcuttur (örneğin, Howell, 2013).

ÖZET

Veri analizinin üç farklı fakat bağlantılı aşaması vardır: Verileri tanımaya başlamak, verileri özetleme ve verilerin anlamının onaylanması. İlk aşamada veriye aşina olmak, veriyi dikkatlice incelemek ve hatalar ile anormal değerler için veriyi kontrol etmek isteriz. Diğer değerler ile birlikte hareket etmiyor gibi görünen aşırı değerlerin, uç değerlerin varlığına duyarlı olmak isteriz. Gövde ve yaprak gösterimini yaratmak, uç değerleri belirlemek amacıyla bir veri kümesindeki sayıların dağılımını görselleştirmenin güzel bir yoludur. İkinci aşamada, betimsel istatistikleri ve grafiksel gösterimleri kullanarak veri kümesini özetlemek isteriz. Merkezî eğilim ölçüleri (ortalama, ortanca, mod) ve yayılım, başka bir deyişle değişkenlik ölçüleri (aralık ve standart sapma) bu noktada özellikle kullanışlıdır. Çalışma, bir bağımsız değişkenin bir bağımlı değişken üzerindeki etkisini içeriyorsa, bağımsız değişkenin bağımlı değişken üzerindeki "etkisinin ne kadar olduğunu" tanımlamak önemlidir. Farklı pek çok çalışmadaki belirli bir değişkenin etkisini özetleyen meta-analizleri yaparken, etki büyüklüğü ölçüleri önemlidir. İki ortalamanın karşılaştırılmasında kullanılan önemli bir etki büyüklüğü ölçüsü, Cohen d'dir.

Veri analizinin üçüncü aşamasında, verilerin anlamının onaylandığı aşamada, çalışmamızda elde ettiğimiz kanıttan hareketle, makul olarak öne ne sürebileceğimizi belirlemek isteriz. Analizin bu aşamasında iki tamamlayıcı yaklaşım kullanılmaktadır: Sıfır hipotezi anlamlılık testi (SHAT) ve güven aralıklarının yapılandırılması. Her iki yaklaşım da, evrendeğerlerin gerçekte ne olduğu hakkında karar vermek amacıyla araştırmacıya yardımcı olmak için, örnekleme değişkenliğinin kestirimine dayanır. Her ne kadar seçkisiz örneklemin ortalaması, evren ortalamasının iyi bir tahmin noktası olsa da, tesadüfî faktörler yüzünden bu tahminin örneklemden örnekleme bir değişkenlik ("hata") göstermesi olasılığı vardır. Ortalamanın tahminî standart hatası,

bir örneklem ortalamasının, evren ortalamasını ne kadar iyi tahmin edebildiğini belirler. SHAT araştırmacının, elde edilen sonuçların "şans eseri" ortaya çıkma olasılığına odaklanmasını sağlar. Güven aralığı, evren değerini (örneğin, evren ortalamasını) içerme olasılığı (genellikle %95) olan değerlerin aralığını belirler. Güven aralıkları, medyada tarama tipi araştırmaların sonuçlarına ilişkin haberlerde rastlamış olabileceğiniz "hata payı" kavramı ile doğrudan ilişkilidir. Aralık daraldıkça, evren değerine ilişkin tahminimiz daha iyi hâle gelir; örneklem büyüklüğünün artması da aralığın tahmin yeteneğini artırır.

İki ortalama arasındaki farka dair güven aralıkları, araştırmadaki iki örneklemin ortalaması ile bunların temsil ettikleri evrenin ortalaması arasındaki farka ait kanıt teşkil eder. Aralığın genişliği, bağımsız değişkenin olası etki büyüklüğüne dair bilgi verir. İki ortalama arasındaki fark için güven aralıkları yapılandırıldığında, bu aralık sıfırı da kapsıyorsa, herhangi bir etkinin olduğunu söylemek istemeyiz. Başka bir deyişle, aralığın içinde sıfır varsa, değişkenin etkisi ile ilgili belirsizlik olduğunu kabul etmeliyiz. Güven aralıkları hem bağımsız gruplar, hem de tekrarlı ölçümler deseni için yapılandırılabilir.

Üç ve daha fazla ortalama varsa, her ortalama için güven aralıkları hesaplanır. Çok sayıda gruptan oluşan bir çalışmada ortalamalar arası farklar ile ilgili olarak sonuca varmak için aralıkların çakışıp çakışmadığına bakılmaktadır. Aralıklar örtüşmediği zaman, örneklem ortalamalarından hareketle tahmin edilen evren ortalamalarının gerçekte farklılaşacağına güvenebiliriz. Bununla birlikte, eğer aralıklar örtüşüyorsa, evren ortalamaları arasında fark olmadığını söylemeyiz; daha ziyade, gerçekte farkın varlığına dair belirsizliği kabul etmeli ve konu ile ilgili daha fazla araştırma yapılmasını beklemeliyiz.

Araştırmacının amacı kestirim yapmaksa (örneğin sınav performansını, sınav kaygısının kâğıt-kalem ile ölçümünden hareketle tahmin etmekte olduğu gibi) genellikle korelasyonel bir çalışma yapılır. Aynı kişi, olay ya da başka bir şeye ait iki faklı ölçüm birlikte değişiyorsa korelasyonun varlığı söz konusudur. Ortalamalar arası fark içeren bir çalışmada yaptığımız gibi, korelasyonel bir çalışmadan elde edilen verileri de dikkatle incelemeli ve özetlemeliyiz. Dağılım grafiği iki puan kümesi arasındaki ilişkiyi betimler; korelasyon katsayısı, dağılım grafiğinde gözlenen ilişkinin niceliksel bir özetini verir. Daha açık bir ifadeyle, korelasyon katsayısı, dağılım grafiğindeki verinin düz bir çizgiye ne kadar uyduğunu gösterir. Bir korelasyon katsayısının değeri –1.00 ile +1.00 arasında değişebilir. Korelasyon katsayısının işareti (– ya da +) ilişkinin yönünü gösterir; katsayının mutlak değeri (0.0 ile 1.00 arasında) ise ilişkinin büyüklüğünü gösterir. Korelasyon katsayısı 1.00'a yaklaştıkça (pozitif veya negatif), dağılım grafiğindeki noktalar düz bir çizgiye yaklaşır ve ilişki daha güçlenir.

Ölçümlerden birindeki değer arttıkça, diğer ölçümdeki değerler de artıyorsa pozitif korelasyon söz konusudur. Negatif korelasyonda ise, ölçümlerden birindeki değerler arttıkça, ikinci ölçümdeki değerler azalır. İki ölçüm arasında bir ilişki (korelasyon) olduğunu biliyor olmak araştırmacıya, ölçümlerden birindeki puanlara bakarak ikinci ölçümdeki puanları kestirme olanağı sağlar. Korelasyon katsayısı 1.00'a veya –1.00'a yaklaştıkça tahmin gücü artar. Korelasyon

katsayısının tek başına, değişkenler arasında nedensel bir ilişki göstermediğini unutmamalıyız.

TEMEL KAVRAMLAR

veri analiz aşamaları 343
  verileri tanımaya başlamak 343
  verileri özetlemek 343
  verilerin anlamını onaylamak 343
gövde ve yaprak gösterimi 348
merkezî eğilim ölçüleri 351
  mod 351
  ortanca 351
  ortalama 352
yayılım ölçüleri (değişkenlik) 352
  aralık 352
  standart sapma 352
  ortalamanın standart hatası 353
  ortalamanın tahminî standart hatası 353
evrendeğer güven aralığı 356
dağılım grafiği 369
doğrusal eğilim 370
negatif korelasyon 371
pozitif korelasyon 371

GÖZDEN GEÇİRME SORULARI

1. Veri analizinin üç temel aşamasını tanımlayın ve her aşamada araştırmacının tipik olarak yapacağı şeyleri belirtin.
2. Bir araştırmacı, çalışmanın bulgularını tamamlamak üzere "analiz öyküsünü" yapılandırırken ne yapmaya çalışmaktadır?
3. Bir araştırmacı, çalışmanın bulgularını analiz etmek üzere bilgisayar yazılımı kullanabilmek için, araştırma yöntemlerine ve istatistiksel işlemlere ilişkin iyi düzeyde bilgiye neden sahip olmalıdır?
4. Şu sayılar için bir gövde ve yaprak gösterimi hazırlayın; sonra veriyi bu şekliyle inceleyerek ne öğrendiğinizi rapor edin. 36, 42, 25, 26, 26, 21, 22, 43, 40, 69, 21, 21, 23, 31, 32, 32, 34, 37, 37, 38, 43, 20, 21, 24, 23, 42, 24, 21, 27, 29, 34, 30, 41, 25, 28
5. Şu veri kümesi için ortalama, ortanca ve modu hesaplayın: 7, 7, 2, 4, 2, 4, 5, 6, 4, 5. Üç merkezî eğilim ölçüsü olan ortalama, ortanca ve modun avantajlarını ve dezavantajlarını açıklayın.
6. Soru 5'teki veri kümesinin standart sapması 1.78'dir. Bu değer size ne ifade ediyor?
7. Ortalamanın tahminî standart hatası, örneklem ortalaması hakkında ne ifade eder?
8. Bellek performansını arttırdığı öne sürülen yeni bir ilacın etkisini araştırmak için, sıçanlar üzerinde bir çalışma yapılmıştır. Bağımlı değişken, bellek ilacı veya plasebo enjekte edildikten sonra, bir labirenti öğrenirken sıçanların yaptıkları hata sayısıdır. Sıçanlar, bellek geliştiren ilaç veya plasebo (kontrol) grubuna seçkisiz olarak atanmışlardır. Her grupta 15 sıçan olmak üzere, toplamda 30 sıçan test edilmiştir. İlaç grubunun ortalaması (ve standart sapması) 11.7 (4.7), kontrol grubunun ise 15.1 (5.1)'dir (Düşük sayılar, daha yüksek performans anlamına geliyor, çünkü labirenti öğrenirken daha az hata yapılmıştır.). Bu çalışmadaki etki büyüklüğü nedir?
9. Neden güven aralığı aynı zamanda "hata payı" olarak adlandırılmaktadır?
10. Seçkisiz oluşturulan 25 öğrencilik bir gruba, üniversitenin yemek hizmeti hakkındaki görüşleri sorulmuştur. Öğrenciler 7'li bir ölçek (1 = berbat 7 = müthiş) üzerinde

görüşlerini belirtmişlerdir. Yirmi beş öğrencinin verdiği ortalama puan 4.7, standart sapması (s) ise 1.2'dir.

A Evren ortalaması için %95'lik güven aralığı nedir?

B Güven aralığının evren ortalamasına ilişkin olarak size ne söylediğini sözcüklerle ifade edin.

11. Soru 8'de verilen iki ortalama arasındaki fark için %95'lik güven aralığı nedir? Bu aralığın yorumu doğru bir şekilde nasıl yapılır?

12. Üç veya daha fazla ortalamanın var olduğu bir çalışmada ortalamalar arası farka ilişkin sonuca varmak için güven aralıklarını nasıl kullanırsınız?

13. Dağılım grafiğinde gösterilen veriyi incelerken, veride doğrusal bir eğilim olup olmadığına bakmak neden önemlidir.

14. Bir araştırmacı, sözcük dağarcığı ile metin kavrama performansı arasında ilişki olup olmadığını araştırmıştır. Altıncı sınıfta olan 15 öğrenciye hem sözcük dağarcığı testi, hem de okunanı anlama testi vermiştir (her iki test de "doğru yüzdesi" cinsinden puanlanmıştır). On beş öğrencinin puanları şöyledir (ilk puanlar sözcük dağarcığı puanlarıdır): 44,67; 24,33; 67,45; 75,54; 34,45; 88,79; 57,67; 44,32; 87,95; 77,67; 87,78; 54,67; 90,78; 36,55; 79,91. Bu puanlar için bir dağılım grafiği çizin ve korelasyon katsayısını hesaplayın.

15. On dördüncü soruda hesapladığınız korelasyon katsayısına dayanarak, sözcük dağarcığını artırmanın, okunanı anlama becerisini arttıracağını iddia edebilir misiniz?.

DÜŞÜNDÜRÜCÜ SORULAR

1 Bir bilişsel psikolog, Gettsyburg savaşını anlatan uzun bir metnin dört farklı sunum şeklinin, akılda tutma düzeyine etkisini araştırıyor. Sunum koşullarına A, B, C ve D diyelim. Altmış dört (N = 64) üniversite öğrencisi bu koşulların her birisine seçkisiz olarak eşit sayıda atanıyor (n = 16). Öğrenciler metni sesli olarak bir defa duyduktan sonra bellek testi uygulanıyor. Bağımlı değişken, yazılı olarak anımsanabilen düşünce sayısıdır. Dört sunum koşulunun her birisinde anımsanan düşünce sayılarının ortalamaları ve standart sapmaları aşağıdaki gibidir:

	A	B	C	D
ort.	16.4	29.9	24.6	19.5
S	4.6	7.1	5.9	6.3

A Dört örneklem ortalamasına dayanarak tahmin edilebilecek evren ortalamalarını, %95 güven aralığına göre hesaplayın.

B Çeşitli evren ortalamaları arasındaki farklar hakkında nasıl bir sonuca varabileceğinizi açıklayarak, güven aralıkları örüntüsünü hesaplayın.

2 Bir gelişim psikoloğu, annelerin çocuğu taşıma davranışının, çocukların uyku örüntüleri üzerindeki etkisini araştırmıştır. Araştırmacı 40 yeni anneden yardım istemiştir. Annelerden 20'sini, yeni doğanın kafası annenin göğsüne temas edecek şekilde taşımaları için eğitmiş, diğer 20 anneye nasıl taşıyacakları konusunda herhangi bir yönerge vermemiştir. Tüm annelerin her 24 saatin sonunda yeni doğanın kaç saat uyuduğunu kaydetmeleri istenmiş ve bu işlem üç ay sürmüştür. Bebeklerin 24 saatlik uyku süreleri, yönerge verilen grup için ortalama 12.6 (S = 5.1); yönerge verilmeyen grup için ortalama 10.1'dir (S = 6.3).

A İki ortalama arasındaki farka ait %95'lik güven aralığını hesaplayın.

B Bu deneyde güven aralıklarından hareketle, eğitimin etkisi hakkında ne söylenebilir?

C Bu deneydeki etki büyüklüğü nedir? Etki büyüklüğü ölçüsünü, Cohen'in küçük, orta ve büyük etki büyüklükleri için olan yönergesini temel alarak, yorumlayın.

3 Bir araştırmacı, üniversite öğrencilerinden zor bir video oyununu, klasik müzik veya hip-hop dinlerken oynamalarını istemiştir. Deneydeki 10 öğrencinin her biri 15 dakika boyunca video

oyununu her iki müzik koşulunda da oynamıştır. Öğrencilerin yarısı önce klasik müzik, sonra hip-hop dinlerken, diğer yarısı müzik türlerini bunun tersi bir sırada dinleyerek oyunu oynamıştır (tekrarlı ölçümler deseninde dengeleme hakkında daha fazla bilgi için bakınız, Bölüm 7). Bağımlı

Öğrenci	Klasik	Hip-hop
1	46	76
2	67	69
3	55	51
4	63	78
5	49	66
6	76	67
7	58	63
8	75	75
9	69	78
10	77	85

değişken, 15 dakikalık oyun süresi içinde yapılan doğru hamle sayısıdır. On öğrenciye ait puanlar aşağıdadır:

A. Her koşul için ortalamaları hesaplayın. Ortalamaları karşılaştırırken nasıl bir eğilim görüyorsunuz?

B. Fark puanlarına ait tahminî standart hatayı hesaplayın.

C. Bu tekrarlı ölçümler desenindeki iki ortalama arasındaki farka ait %95'lik güven aralığını hesaplayın.

D Bu bulgulara ilişkin analiz sonucunda, müzik türünün performans üzerindeki etkisi ile ilgili bir sonuca varınız.

4 Bir sosyal psikolog, kağıt-kalemle ölçülmüş ön yargı düzeyi ile, ırksal profillemenin bir suç caydırıcısı olarak kullanılmasına yönelik tutumlar arasındaki ilişkiyi belirlemek istemiştir. Bu amaçla, dönemin başında genel psikoloji dersindeki öğrencilerden altı farklı anketi doldurmaları istenmiştir. Anketler, ön yargının ölçümü için verilmiştir. Dönemin ilerleyen zamanlarında öğrenciler, suçlu davranışı ve yasalara uymayı sağlamaya dair taktikler ile ilgili bir deneye katılmak üzere davet edilmişlerdir. Deneyin bir parçası olarak öğrenciler, ırksal profillemenin suç caydırıcısı olarak kullanılmasıyla ilgili sorular içeren bir anketi yanıtlamışlardır. Araştırmacı, önceden alınan ön yargı ölçümünün, kişilerin ırksal tutumlarını kestirip kestiremeyeceğini bulmak istemiştir. Ön yargı ölçümündeki daha yüksek puanlar daha fazla ön yargıya karşılık geliyor; ırksal tutum ölçeğindeki daha yüksek puanlar ise ırksal belirlemeye daha fazla destek vermek anlamına geliyor. İki ölçüm için 22 öğrenciden elde edilen puanlar aşağıdadır:

Öğrenci	1	2	3	4	5	6	7	8	9	10	11
Ön yargı	19	15	22	12	9	19	16	21	24	13	10
Profilleme	7	6	9	6	4	7	8	9	5	5	7

Öğrenci	12	13	14	15	16	17	18	19	20	21	22
Ön yargı	12	17	23	19	23	18	11	10	19	24	22
Profilleme	4	8	9	10	10	5	6	4	8	8	7

A Bu iki ölçüm arasındaki ilişkiyi gösteren bir dağılım grafiği çizin.

B Dağılım grafiğini inceleyin ve veride doğrusal bir eğilimin olup olmadığına karar verin.

C Bu veriler için korelasyon katsayısını hesaplayın ve ilişkinin yönü ile gücünü yorumlayın.

D Korelasyon analizine dayanarak araştırmacı şu sonuca varır: Ön yargılı düşünce insanların, güvenlik güçlerinin ırkçı davranışlarını desteklemelerine neden olur. Varılan bu sonucu, korelasyonel kanıtların niteliğine ilişkin bilgilerinize dayanarak yorumlayınız.

İleri Alıştırmaların Cevapları

İfade 1 ve 5 doğru; 2, 3 ve 4 yanlış.

Düşündürücü Soruların Cevapları

A Dört grup için $s_{birleşik}$'i hesaplamakla işe başlayın, her grup için standart sapmanın elde edilebildiğinden ve $s_{birleşik}$'in formülünde varyansların kullanıldığından emin olun. Bu nedenle $n - 1$'e bölünmeden önce her bir standart sapmanın karesi alınmalıdır. $s_{birleşik}$'in değeri 6.04'tür. Buradan hareketle, ortalamanın tahminî standart hatası ($s_{\bar{X}}$) $6.04/\sqrt{n}$, yani 1.51'dir. Tablo A.2 incelendiğinde, .05 düzeyi için kritik t değeri 2.00 olarak bulunur (serbestlik derecesi 60). Ortalamalar için güven aralıkları ise şu şekildedir:

A $16.4 \pm (2.00)(1.51) = 13.38$ ile 19.42 arasında

B $29.9 \pm (2.00)(1.51) = 26.88$ ile 32.9 arasında

C 24.6 ± (2.00)(1.51) = 21.58 ile 27.62 arasında

D 19.5 ± (2.00)(1.51) = 16.48 ile 22.52 arasında

B (*İpucu*: her bir grubun ortalama performansını ve güven aralıkları için ortalama etrafındaki çubukları, kolonlarda gösteren bir şekil çizmek yararlı olabilir. Kutu 11.5'teki bilgileri de gözden geçirebilirsiniz.) A aralığının sadece D aralığı ile örtüştüğü görülmektedir. C ve D aralıkları da örtüşmektedir. Gözlenen grup ortalamaları, evren değerlerinin bulunduğu bölge ile ilgili en iyi tahminimiz olsa da, güven aralıkları da ayrıca yaptığımız tahminin kesinliği hakkında bilgi vermektedir. Bu verilerden hareketle, örneklem ortalaması A tarafından tahmin edilen evren ortalamasının, B ve C tarafından temsil edilen evren ortalamalarından farklılaştığı sonucuna varabiliriz. A ve D arasındaki farka ilişkin ise kesin bir yargıda bulunmaktan kaçınmalıyız. Ayrıca B ve D evren ortalamaları arasında fark olduğu, fakat B ve C arasındaki gerçek farka ilişkin kesin bir şey söyleyemeyeceğimiz sonucunu da çıkarabiliriz.

BÖLÜM ON İKİ

Veri Analizi ve Analizin Yorumlanması: Kısım 2. İstatistiksel Anlamlılığın Test Edilmesi ve Analizin Öyküsü

BÖLÜMÜN ANA HATLARI

GENEL BAKIŞ
SIFIR HİPOTEZİ ANLAMLILIK TESTİ (SHAT)
DENEYSEL DUYARLILIK VE İSTATİSTİKSEL GÜÇ
SHAT: İKİ ORTALAMANIN KARŞILAŞTIRILMASI
Bağımsız Gruplar
Tekrarlı Ölçümler Desenleri
İSTATİSTİKSEL ANLAMLILIK VE BİLİMSEL YA DA PRATİK ANLAMLILIK
İKİ ORTALAMANIN KARŞILAŞTIRILMASI İLE İLGİLİ ÖNERİLER
İKİ ORTALAMAYI KARŞILAŞTIRIRKEN SONUÇLARIN RAPOR EDİLMESİ
İKİDEN FAZLA KOŞUL İÇİN VERİ ANALİZİ
TEK FAKTÖRLÜ BAĞIMSIZ GRUPLAR DESENİ İÇİN ANOVA
Üç veya Daha Fazla Bağımsız Grup İçeren Desenler İçin Etki Büyüklüğünün Hesaplanması
Bağımsız Gruplar Deseni İçin Gücün Değerlendirilmesi
Çoklu Grup Deneylerinde Ortalamaların Karşılaştırılması
TEKRARLI ÖLÇÜMLER VARYANS ANALİZİ
BAĞIMSIZ GRUPLAR DESENLERİ İÇİN İKİ FAKTÖRLÜ VARYANS ANALİZİ
Etkileşimin Olduğu Karmaşık Desen Analizi
Etkileşimin Olmadığı Durumda Analiz
Bağımsız Grupları İçeren İki Faktörlü Desene Ait Etki Büyüklükleri
KARMAŞIK DESENLERİN ANALİZİNDE GÜVEN ARALIKLARININ ROLÜ
KARMA DESEN İÇİN İKİ FAKTÖRLÜ VARYANS ANALİZİ
KARMAŞIK DESEN İÇİN SONUÇLARIN RAPOR EDİLMESİ
ÖZET

GENEL BAKIŞ

Bölüm 11'de veri analizinin üç aşamasına değindik: *Verileri tanımaya başlama, verileri özetleme* ve *verilerin anlamını onaylama.* Veri analizinin son aşamasında, davranışa ilişkin bir iddiada bulunmak için yeterli kanıtımızın olup olmadığını değerlendireceğiz. Elde ettiğimiz verilerden hareketle, davranışa ilişkin ne söyleyebiliriz? Bu aşama bazen *onaylayıcı veri analizi (confirmatory data analysis)* olarak da adlandırılır (örneğin, Tukey, 1977). Bu noktada verinin bize ne söylediğini onaylamaya çalışırız. Bölüm 11'de verinin bize ne söylediğini onaylamak için güven aralıklarının kullanımı üzerinde durduk. Bu bölümde ise istatistiksel anlamlılığın test edilmesine, daha iyi bilinen adıyla, *sıfır hipotezi anlamlılık testine* (SHAT) odaklanarak, onaylayıcı veri analizi üzerindeki tartışmamıza devam edeceğiz.

SHAT, psikolojide onaylayıcı veri analizi için kullanılan en yaygın yaklaşımdır. Bununla birlikte, istatistiksel anlamlılığın test edilmesi, sürekli olarak haklı eleştiriler almıştır (örneğin; Cohen, 1995; Hunter, 1997; Loftus, 1991, 1996; Meehl, 1967; Schmidt, 1996). Araştırmacılar onlarca yıldır anlamlılık testlerini yanlış kullanmış (ve yanlış yorumlamış) ve uyarıları göz ardı etmişlerdir (örneğin, Finch, Thomason, ve Cumming, 2002). SHAT'yi tamamen gözden çıkarmamız gerektiği öne sürülmüştür (örneğin, Hunter, 1997; Schmidt, 1996). Bununla birlikte uzmanların çoğu SHAT'yi kullanmaya devam etmemiz, fakat kullanımına ilişkin olarak dikkatli olmamız gerektiğini önermektedir (örneğin; Abelson, 1995, 1997; Chow, 1988; Estes, 1997; Greenwald, Gonzalez, Harris ve Guthrie, 1996; Hagen, 1997; Howell, 2013; Krueger, 2001; Mulaik, Raju ve Harshman, 1997). Psikoloji camiası içindeki bu tartışmanın sonucu ne olursa olsun, konu ile ilgili neredeyse evrensel olan bir görüş birliği bulunmaktadır: (a) SHAT'nin neyi yapabildiğini ve yapamadığını tam olarak anlamak, ve (b) veri analizi ile ilgili alternatif yöntemlerin kullanımını, özellikle de güven aralıklarının ve etki büyüklüklerinin kullanımını artırmak. Bazen bu alternatif teknikler SHAT'nin yerine geçmekte, bazen de SHAT'yi tamamlamaktadır.

Bu bölümde, öncelikle SHAT'ye ilişkin bir gözden geçirme sunacağız. Sonrasında ise deneysel duyarlılık ve istatistiksel güç ile ilgili önemli kavramları tartışacağız. Ayrıca Bölüm 11'de iki ortalama arasındaki farka ait güven aralıklarını hesaplamak için kullandığımız veriyi kullanarak veri analizinde SHAT yaklaşımını göstereceğiz. Aynı veriyi kullanmakla, SHAT ile elde edilen bilgiyi, güven aralıkları ile elde edilen bilgiyle karşılaştırabileceğiz. SHAT'yi temel alarak ne söyleyebileceğimiz ve söyleyemeyeceğimiz üzerinde duracağız ve SHAT ile elde edilen bilginin, güven aralıkları kullanılarak elde edilen bilgiyi tamamlayıcı bir niteliğinin olduğunu belirteceğiz. Son olarak, iki ortalamayı içeren durumlarda, davranışa ait iddiaya yönelik bir kanıtı değerlendirirken takip etmek üzere bazı öneriler sunacağız ve çalışmanız için bir analiz öyküsünün nasıl yaratılacağını göstereceğiz.

İkiden fazla grup içeren çalışmalarla ilgili onaylayıcı veri analizi için en yaygın şekilde kullanılan teknik bir tür SHAT'dir ve *varyans analizi (ANOVA)* adı verilir. ANOVA'nın kullanılmasındaki mantık, ANOVA ile ilgili hesaplamalar ve ANOVA sonuçlarının yorumlanması, bu bölümün ikinci yarısında tartışılacaktır.

SIFIR HİPOTEZİ ANLAMLILIK TESTİ (SHAT)

- Sıfır hipotezi testi, bir deneyin grupları arasındaki ortalama farkların, basitçe hata değişkenliğinden kaynaklı olarak beklenen farklardan büyük olup olmadığını belirlemek için kullanılır.

- Sıfır hipotezi testindeki ilk adım, grupların farklılaşmadığını, yani bağımsız değişkenin bir etkisinin olmayacağını varsaymaktır (sıfır hipotezi).
- Sıfır hipotezinin doğru olduğunu varsayarak, deneyin gözlenen sonucunun ortaya çıkma olasılığını tahmin etmek için olasılık kuramından yararlanılır.
- İstatistiksel olarak anlamlı bir sonuç, eğer sıfır hipotezi doğru olsaydı, meydana gelme olasılığı çok küçük olan bir sonuçtur.
- Bir deneyin sonucu hakkında verilen kararlar olasılığa dayandığı için, bu kararlar verilirken hata yapılabilir; bunlar Tip I (doğru bir sıfır hipotezinin reddedilmesi) veya Tip II (yanlış bir sıfır hipotezinin reddedilmemesi) şeklinde olabilir.

İstatistiksel çıkarım hem tümevarımsal, hem de dolaylıdır. Tümevarımsaldır, çünkü güven aralıklarını hesaplarken yaptığımız gibi, deneylerimizde test ettiğimiz belirli örneklemlerden hareketle evrene ilişkin genel sonuçlar çıkarırız. Bununla birlikte, güven aralıklarının kullanıldığı yaklaşımdan farklı olarak, bu türdeki istatistiksel çıkarım ayrıca dolaylıdır çünkü işe sıfır hipoteziyle ilgili varsayımla başlanmaktadır. **Sıfır hipotezi (H_0)**, bağımsız değişkenin bir etkisinin olmadığını öne süren varsayımdır. Bu varsayımı yaptığımızda, deneyimizde elde ettiğimiz farkı (veya daha büyük bir farkı), sıfır hipotezinin doğru olması durumunda elde etme olasılığını belirlemek için olasılık kuramından yararlanabiliriz. Eğer bu olasılık küçük ise sıfır hipotezini reddederiz ve bağımsız değişkenin bağımlı değişken üzerinde bir etkisinin olduğu sonucunu çıkarırız. Sıfır hipotezini reddetmemize neden olan sonuçlar *istatistiksel olarak anlamlı*, şeklinde ifade edilir. İstatistiksel olarak anlamlı bir sonuç, deneyde elde ettiğimiz farkın, tek başına hata değişkenliğinin (yani şans eseri ortaya çıkan değişkenliğin) sorumlu olduğu farktan büyük olduğu anlamına gelir (bk. Kutu 12.1).

Temel Kavram

İstatistiksel olarak anlamlı bir sonuç, sıfır hipotezinin doğru olması durumunda, ortaya çıkma olasılığı çok düşük olacak olan bir sonuçtur. Peki, bunun için olasılığın ne kadar küçük olması yeterlidir? Her ne kadar bu önemli soru için kesin bir yanıt bulunmasa da bilim insanları arasındaki ortak görüş, eğer sıfır hipotezi doğru olsaydı, her 100 sonuçtan 5'ine karşılık gelen olasılığa sahip sonuçların (ya da .05) istatistiksel olarak anlamlı olduğudur. Bir sonucun istatistiksel olarak anlamlı olduğuna karar vermek için kullandığımız olasılığa **anlamlılık düzeyi** denir. Anlamlılık düzeyi, Yunan harfi olan alfa (α) ile gösterilir. Bu nedenle, bahsettiğimiz .05 düzeyindeki anlamlılığı, $\alpha = .05$ olarak bildiririz.

Temel Kavram

Sonuçlarımızın anlamlı olması ne demektir? Bu durumda elde ettiğimiz en kullanışlı bilgi, ilgi çekici bazı şeylerin gerçekleşmiş olduğudur. Daha açık bir ifadeyle, gözlenen sonuca ilişkin olasılık küçüldükçe, deney tekrarlandığı zaman (replication) istatistiksel olarak anlamlı bir sonucu tekrar bulma olasılığı artmaktadır. Fakat bu ifade ile ne kastettiğimize dikkat etmeliyiz. Araştırmacılar bazen yanlış bir biçimde, $p < .05$ şeklindeki bir olasılığı olan bir sonucu, "Eğer çalışma tekrarlanırsa bu sonuç 95/100 kez elde edilecektir" biçiminde yorumlamaktadır. Bu yorum, en basit ifadeyle, doğru değildir. İstatistiksel anlamlılığı elde etmek (yani $p < .05$), sonuçların tekrarlanma olasılığı hakkında bir şey söylemez. Örneğin, .05 olasılığın altındaki bir sonuç (istatistiksel olarak anlamlı), araştırma tam olarak tekrar edilseydi, istatistiksel olarak anlamlı olma olasılığı yaklaşık olarak sadece 50:50 oranında olacaktı (Greenwald ve ark., 1996). Öte yandan, sonuçlara ilişkin tam olasılığı biliyor olmak, deney tekrarlansaydı neler olabileceğine dair bilgi verir. Önceki bir bulguya ilişkin olasılık

KUTU 12.1

YAZI MI TURA MI? YAZI TURA ATMAK VE SIFIR HİPOTEZİ

Şu ikilem üzerinde düşünürseniz, muhtemelen istatistiksel çıkarım sürecinin önemini kavrayacaksınız. Bir arkadaşınız muzip bir gülümsemeyle, az önce lokantada yediğiniz yemeğin hesabını kimin ödeyeceğini belirlemek için yazı tura atmayı teklif ediyor olsun. Arkadaşınızın elinde yazı tura atmaya hazır bir parası var. Bu durumda arkadaşınızın parasının hileli olup olmadığını doğrudan test etmek (yani, parayı kontrol etmeyi istemek) sizin için iyi olabilir. Fakat arkadaşınıza güvensizlik anlamına geleceği için bunu tercih etmiyorsunuz. Böyle bir durumda yapabileceğiniz en iyi şey, paranın hileli olmadığını varsaymak ve beklenen 50:50 yazı tura oranından sistemli şekilde sapan sonuçlar çıkıp çıkmayacağını görerek, parayı dolaylı yoldan test etmektir. Eğer para beklenen 50:50 oranını vermezse (çok sayıda denemeden sonra), arkadaşınızın el altından size yemek parasını ödetmeye çalıştığı sonucunu çıkarmanız olasıdır. Benzer şekilde, deneylerimizde elde edilen bir sonuç için de istatistiksel anlamlılığa yönelik olarak doğrudan bir test yapmak isteriz. Ne var ki, yapabileceğimiz en iyi şey, elde ettiğimiz sonucu, yazı ve turanın gözlenme sıklığı arasında fark olmadığına ilişkin beklenen sonuçla karşılaştırmaktır. *Sıfır hipotezi testini anlamanın temelinde şu kabul yatar: Bir sonucun tek nedeninin tümüyle şans faktörü olduğunu varsaymak koşuluyla, bu sonucun gerçekleşme olasılığını tahmin etmek için olasılık yasalarını kullanabiliriz.* Bu, sonuca varmak için arkadaşınızın parasını birçok kez atmanızdan farklı değildir. Tamamen şansa bağlı olduğunda, denemelerin %50'sinde tura, %50'sinde ise yazı geleceğini biliyorsunuz. Para birçok kez atıldığında, bu olası sonuçtan farklı olan herhangi bir sonuç, sizi şans faktöründen farklı başka bir şeyin etkili olduğu, yani arkadaşınızın parasının hileli olduğu sonucuna varmaya itecektir.

küçüldükçe, araştırmanın sonraki bir tam tekrarında (replication), istatistiksel olarak anlamlı ($p < .05$) bir sonuç elde etme olasılığı artacaktır (örneğin, Posavac, 2002). Sonuç olarak, Amerikan Psikoloji Birliği'nin (APA) de önerdiği gibi, *SHAT gerçekleştirirken her zaman sonuçlara ilişkin kesin olasılığı bildirin.*

Katı bir yorum yaparsak, bir çıkarımsal istatistiksel test yaparken elde edilebilecek sadece iki olası sonuç vardır: Ya sıfır hipotezi *reddedilir* ya da sıfır hipotezi *reddedilmez.* Burada sıfır hipotezini kabul etmenin de bir alternatif olduğunu *söylemediğimize* dikkat edin. Şimdi bu duruma açıklık getirelim.

Yaptığımız bir deneyde bağımsız değişkenin etkisinin istatistiksel olarak anlamlı olmadığını gözlediğimizde, sıfır hipotezini reddetmeyiz. Ama farkın olmadığına dair sıfır hipotezini kabul etmek zorunda da değiliz. Deneyimizde, bağımsız değişkenin etkisini gözlememizi engellemiş olan bazı faktörler de olmuş olabilir (örneğin; katılımcılara verilen yönergelerin belirsiz olması, bağımsız değişkenin gerektiği gibi uygulanmaması gibi). Daha sonra göstereceğimiz gibi, sıfır hipotezinin reddedilmemesinin temel bir nedeni, örneklemin çok küçük olması durumudur. Sıfır hipotezinin doğru olduğunu kanıtlamanın mantıksal olarak olanaksız olduğunu bilsek de elimizde araştırma için hangi bağımsız değişkenlerin peşine düşmeye değmeyeceğini belirlemeye yarayacak bazı yöntemler bulunmalıdır. Bu tür bir kararın verilmesinde SHAT yardımcı olmaktadır. İstatistiksel olarak sonucun anlamlı çıkmaması şu anlama gelir: Bağımsız değişkenin, davranışı tesadüfî ve önemsiz bir nedenden daha fazla etkilediği sonucuna varmakta temkinli olmalıyız. Bu noktada daha fazla bilgiye gereksinim

TABLO 12.1 ÇIKARIMSAL İSTATİSTİKLER KULLANILARAK VERİLEN KARARLARIN OLASI SONUÇLARI

	Olası Durumlar	
	Sıfır Hipotezi Yanlış	Sıfır Hipotezi Doğru
Sıfır hipotezinin reddi	Doğru karar	Tip I hata
Sıfır hipotezinin reddedilememesi	Tip II hata	Doğru karar

duyacaksınız; örneğin, örneklem büyüklüğü ve etki büyüklüğünü ele almak gibi (sonraki "Deneysel Duyarlılık ve İstatistiksel Güç" bölümüne bakın).

İstatistiksel çıkarımla ve kararlarımızı verirken olasılıkları temel almamızla ilgili sıkıntılı bir durum var. Hangi karara varırsanız varın ve bu karara ulaşırken ne kadar dikkatli davranırsanız davranın, her zaman bir hata yapma olasılığınız vardır. Araştırmacının varabileceği iki olası karar ve iki olası durum Tablo 12.1'de görülmektedir. İki olası durum, bağımsız değişkenin davranış üzerinde etkisinin olduğu ve olmadığı durumlardır. Araştırmacının verebileceği iki olası doğru karar, tablonun sol üst ve sağ alt hücrelerinde gösterilmiştir. Bağımsız değişkenin bir etkisi varsa, araştırmacı sıfır hipotezini reddetmelidir; eğer bağımsız değişkenin bir etkisi yoksa araştırmacı sıfır hipotezini reddetmemelidir.

İki potansiyel hata (Tip I hata ve Tip II hata), Tablo 12.1'in diğer hücrelerinde gösterilmiştir. Bu hatalar, istatistiksel çıkarımın nitelik olarak olasılığa dayanmasından kaynaklanır. Bir sonucun sıfır hipotezi altında ortaya çıkma olasılığı .05'ten az olduğu için, o sonucun istatistiksel olarak anlamlı olduğuna karar verdiğimiz zaman şunu kabul ediyoruz: Her 100 testten 5'inde, sıfır hipotezi doğru olsa bile söz konusu sonuç elde edilebilir. Bu nedenle anlamlılık düzeyi, **Tip I hata** yapma olasılığını göstermektedir: Yani, doğru olan sıfır hipotezini reddetmek. Tip I hata yapma olasılığı, basitçe, anlamlılık düzeyi daha tutucu hâle getirilerek, örneğin .01'e çekilerek, düşürülebilir. Bu yaklaşımın sorunu ise, **Tip II hata** yapma olasılığını arttırabilmesidir: Yani, yanlış bir sıfır hipotezinin reddedilememesi.

Temel Kavramlar

Tip I hata ve Tip II hataları ile ilgili sorun bizi durdurmamalıdır, fakat bu sorun araştırmacıların anlamlılık testlerinin kullanıldığı bir deneyin sonuçlarını açıklarken "kanıtlamak" sözcüğünü neden nadiren kullandıklarını anlamamızı sağlar. "Kanıtlamak" sözcüğünü kullanmak yerine araştırmacılar sonuçları, "hipotez ile tutarlı" veya "hipotezi onaylamaktadır" ya da "hipotezi desteklemektedir" şeklinde açıklamaktadırlar. Bu tür kesin olmayan ifadeler, her zaman Tip I ve Tip II türünden hata yapma olasılığının var olduğunu dolaylı olarak kabul etmenin bir yoludur. .05 düzeyindeki anlamlılık, ortak bir denge kurup, her iki hata türünü de birçok kez yapmaktan kaçınmamızı sağlayan bir uzlaşma noktasını temsil etmektedir. Tip I ve Tip II sorunu ayrıca *istatistiksel çıkarımın, deneysel bir sonucun güvenirliğinin test edilmesi için asla replikasyonun yerini alamayacağını,* bize hatırlatmaktadır.

DENEYSEL DUYARLILIK VE İSTATİSTİKSEL GÜÇ

- Duyarlılık, eğer gerçekten bir bağımsız değişkenin etkisi varsa, bir deneyin, o bağımsız değişkenin etkisini gösterebilme olasılığı demektir.
- Güç, bir istatistiksel testin, grup farklılıklarının olmadığına dair sıfır hipotezinin araştırmacı tarafından doğru şekilde reddedilmesini sağlayabilme olasılığı demektir.
- İstatistiksel testlerin duyarlılığı anlamlılık düzeyinden, uygulama etkisinin büyüklüğünden ve örneklemin büyüklüğünden etkilenir.
- Araştırmacılar için istatistiksel gücü artırmanın temel yolu, örneklemi büyütmektir.
- Tekrarlı ölçümler desenlerinin, bağımsız gruplar desenlerine göre daha duyarlı ve istatistiksel gücünün daha yüksek olması muhtemeldir; bunu nedeni, hata değişkenliğine ilişkin tahminlerin tekrarlı ölçümler desenleri için daha küçük olmasıdır.
- Psikolojide, SHAT'nin kullanıldığı araştırmalarda Tip II hata, Tip I hatadan daha yaygındır.
- Sonuçlar istatistiksel olarak anlamlı olmadığı zaman (yani $p > .05$) sıfır hipotezinin doğru olduğu sonucunu çıkarmak yanlıştır.

Temel Kavram

Bir deneyin duyarlılığı, eğer bağımsız değişkenin gerçekten bir etkisi varsa, deneyin, bağımsız değişkenin bu etkisini gösterme olasılığıdır (bk. Bölüm 7). Deneyin duyarlılığının, istatistiksel testin ise gücünün olduğu söylenir. Bir istatistiksel testin **gücü**, yanlış bir sıfır hipotezini reddedebilme olasılığıdır. Sıfır hipotezi, "fark yok" hipotezidir ve eğer bağımsız değişken bir farka neden olmuşsa, bu durumda sıfır hipotezi yanlıştır ve reddedilmelidir. Tip II hatayı, yanlış bir sıfır hipotezini reddedememe olarak tanımladığımızı anımsayın. Güç aynı zamanda, 1 eksi Tip II hata yapma olasılığı olarak tanımlanır.

Güç, var olan bir etkiyi "görme" olasılığımızı bize söyler ve çalışmanın tekrar edilebilirliği ile ilgili bir tür tahmindir. Güç, yanlış olan bir sıfır hipotezinin reddedilme olasılığını verdiği için, bu sayede gerçekten var olan bir etkiyi gözden kaçırma olasılığını bilebiliriz. Örneğin, bir sonuç anlamlı değilse ve güç sadece .30 ise biliriz ki bu özelliklere sahip olan bir çalışma, 10 seferde ancak 3 kez gözlenebilen büyüklüğe eşit bir etkiyi saptayabilir. Bu nedenle, söz konusu çalışmayı 10 kez yaparsak, bunların 7'sinde etkiyi gözden kaçıracağız. Böyle bir durum karşısında, çalışma daha yüksek bir güce sahip olacak şekilde tekrar yapılıncaya kadar, yargıya varmakta acele etmek istemeyiz.

İstatistiksel bir testin gücü, şu üç faktör arasındaki etkileşime bağlıdır: İstatistiksel anlamlılığın düzeyi, uygulamanın etkisinin büyüklüğü ve örneklemin büyüklüğü (Keppel, 1991). Bununla birlikte, pratik nedenlerle *örneklem büyüklüğü, araştırmacının istatistiksel gücü kontrol etmesi için başlıca faktördür.* Farklı büyüklükteki etkilerin saptanabilmesi için gerekli olan örneklem büyüklükleri arasındaki farklar dramatik olabilir. Örneğin Cohen (1998) üç düzeyli bir bağımsız değişkeni olan ve bağımsız gruplar deseni kullanan bir deney için gerekli olan örneklem büyüklüğünü bildirmiştir. Buna göre, uygulamaya ait büyük bir etkiyi saptayabilmek için gerekli olan örneklem büyüklüğü 30, orta düzeyde bir uygulama etkisini saptayabilmek için gerekli olan örneklem büyüklüğü 76 ve küçük bir uygulama etkisini saptayabilmek için gerekli olan

örneklem büyüklüğü 464'tür. Başka bir deyişle, büyük bir etkiye kıyasla, küçük bir etkiyi saptamak için 15 kat fazla katılımcıya ihtiyaç vardır!

Tekrarlı ölçümler kullanan deneyler yapmak da araştırmacının kullandığı istatistiksel analizlerin gücünü etkileyebilir. Bölüm 7'de açıklandığı gibi, genel olarak tekrarlı ölçümleri içeren deneyler, bağımsız grup deneylerine göre daha duyarlıdır. Daha küçük hata değişkenliği, bir deneydeki küçük uygulama etkilerini saptama yeteneğini artırmaktadır ve bu, istatistiksel gücün tanımına karşılık gelmektedir –eğer varsa, küçük uygulama etkilerini saptayabilme becerisi.

SHAT'ye değinirken, Tip I hata olarak adlandırılan şeyin alfaya (burada, .05) eşit olduğundan bahsetmiştik. Mantıksal olarak, bu tür bir hata yapmak için sıfır hipotezinin yanlış olabilme özelliğinin olması gerekir. Aslında çeşitli eleştirilerde, sıfır fark şeklinde tanımlanan sıfır hipotezinin "her zaman yanlış" olduğu (örneğin, Cohen, 1995, s. 1000) veya biraz daha tutucu bir yorum olarak "nadiren doğru" olduğu tartışılmıştır (Hunter, 1997, s. 5). Evren ortalamaları arasında sıfır fark olduğu şeklinde ifade edilen sıfır hipotezinin herhangi bir durumda doğru olup olmadığı tartışma konusudur. Bazıları sıfır hipotezinin her zaman yanlış olduğunu, yani evren ortalamaları arasında her zaman en azından bir miktar fark olduğunu öne sürmektedir. Yine de bir çok durumda, farklılığın olmadığına dair hipotez karşısında bir etkiyi test etmek önemlidir (bk. Abelson, 1997; Mulaik ve ark., 1997). Araştırmacılar, bazı durumlarda eğer bir etki yoksa veya en azından tesadüfî bir etkiden daha fazlası yoksa bir etkinin var olduğu sonucunu *çıkarmamanın* önemli olduğu konusunda uyanık olmalıdır (bk. Kutu 12.2). Tip II hatanın, yani gerçek bir etkinin olmadığını söyleme olasılığının, genellikle .05'ten çok daha büyük olduğu yönünde uzlaşma vardır. Psikoloji ile ilgili birçok çalışmada bu tür bir hata, düşük istatistiksel güç nedeniyle gerçekleşmektedir.

KUTU 12.2

SIFIR HİPOTEZİNİN KABUL EDİLDİĞİ DURUMLAR VAR MIDIR?

Şimdiye kadar aksini söylemiş olsak da araştırmacıların sıfır hipotezini kabul etmeyi tercih ettiği (hipotezi reddedememe yerine) bazı durumlar olabilir. Yeaton ve Sechrest (1986, ss. 836–837), farkın olmadığına dair bulguların, uygulamalı araştırmalar için kritik önemi olduğunu ikna edici şekilde tartışmıştır. Değindikleri şeyi vurgulayabilmek için sordukları birkaç soruyu ele alalım: Kreşe gönderilen çocuklar, evde kalan çocuklara göre entelektüel, sosyal ve duygusal açıdan daha fazla mı ilerleme gösterirler? Diğerlerine göre daha ucuz olan ve daha az yan etkisi olan, yeni geliştirilen bir ilaç kalp krizlerini engellemede, mevcut standart uygulama kadar etkili midir?

Bu tür önemli sorular, sıfır hipotezini kabul etmenin (etkinin olmamasının) kuramsal bir mesele olmaktan öte bir şey olduğunu açıkça göstermektedir; başka bir deyişle, doğru kararı verebilmek bazen bir ölüm kalım meselesidir. Frick (1995), sıfır hipotezini asla kabul etmemenin psikoloji için ne arzulanır, ne de pratik yararı olan bir şey olmadığını savunmuştur. Güvenilir bir biçimde, dikkate değer bir fark olmadığını söylemek isteyebileceğimiz bazı durumlar söz konusu olabilir (ayrıca bk. Shadish, Cook ve Campbell, 2002; Howell, 2013, ss. 91–92).

İstatistiksel güç düşük olduğunda, Tip II hataların oluşması muhtemeldir ve düşük güç, literatürde birçok çalışmada gözlenmiştir: *Psikolojide, SHAT'nin kullanıldığı araştırmalardaki en yaygın hata, Tip II hatadır.* Bunun nedeni basitçe, istatistiksel anlamlılığı elde etmemiş olmamızın, bir etkinin var olmayacağı anlamına gelmemesidir (örneğin, Schmidt, 1996). Aslında, etki büyüklüğünün ölçülmesinin altında yatan önemli nedenlerden biri, etkinin istatistiksel olarak anlamlı olup olmadığından bağımsız olarak, elde edilen etkinin, diğer çalışmalarda elde edilenlerle karşılaştırılabiliyor olmasıdır. Meta-analiz, bu amaçla yapılır (bk. Bölüm 6). Her ne kadar anlamlı olmayan bir bulgu, bize etkinin olmadığını söylemiyor olsa da çalışmamızın yeterli istatistiksel güçle yapıldığını varsayarsak, anlamlı olmayan bir bulgu, etkinin çok küçük olduğunu, yani dikkate almaya değer olmadığını gösterebilmektedir.

Çalışmanızın gücünü çalışmayı yapmadan *önce* belirlemek için, öncelikle deneyinizden beklenen etki büyüklüğünü tahmin etmelisiniz. Önceki çalışmalarda, araştırılan bağımsız değişken için elde edilen etki büyüklüklerini incelemek, tahmininizi yapmanıza yardımcı olacaktır. Etki büyüklüğü tahmin edildikten sonra ise etkiyi "görebilmek" için kullanmanız gereken örneklem büyüklüğü ile ilgili bilgi edinmek amacıyla "güç tablolarını" incelemelisiniz. Güç analizini yapmak için gerekli olan bu adımlar, çeşitli istatistik kitaplarında daha ayrıntılı şekilde açıklanmıştır (örneğin, Zechmeister ve Posavac, 2003) ve güç tabloları da Web'den elde edilebilir. *Test ettiğiniz etki büyüklüğüne ilişkin iyi bir tahmininiz varsa, araştırmayı yapmadan önce güç analizine başvurmanızı ısrarla öneriyoruz.*

Güç tabloları, araştırma yapıldıktan sonra da kullanılır. Çalışma tamamlandığında ve elde edilen bulgular istatistiksel olarak anlamlı çıkmadığında APA *Yayın Kılavuzu* (2010), çalışmanızın gücünün rapor edilmesini önermektedir. Bunu yaparak diğer araştırmacılarla, var olan bir etkinin saptanma olasılığı hakkında iletişim kurmuş olursunuz. Eğer söz konusu olasılık düşükse araştırmacı camiası çalışmanızın daha yüksek güce sahip bir tekrarı yapılana kadar, bulgularınızla ilgili olarak bir yargıya varmak konusunda acele etmek istemeyebilir. Öte yandan, yeterli güce sahip bir çalışmadan elde edilen ve istatistiksel olarak anlamlı olmayan bir sonuca bakarak, araştırma camiası söz konusu etkiyi daha fazla araştırmaya değer bulmayabilir.

SHAT: İKİ ORTALAMANIN KARŞILAŞTIRILMASI

- Farklı gruplardan elde edilen iki ortalamayı karşılaştırırken kullanılacak olan uygun çıkarımsal test, bağımsız gruplar *t*-testidir.
- SHAT kullanıldığı zaman etki büyüklüğü daima rapor edilmelidir.
- Aynı deneklerden (ya da eşleşmiş gruplardan) elde edilen iki ortalamayı karşılaştırmak için uygun olan çıkarımsal istatistik, tekrarlı ölçümler (denek içi) *t*-testidir.

Şimdi, iki ortalama arasındaki farkı karşılaştırmak için SHAT'nin nasıl kullanıldığını göstereceğiz. Öncelikle, iki bağımsız ortalamanın var olduğu bir çalışma ele alınacaktır. Bu çalışmaya ait veriler, Bölüm 11'de açıkladığımız sözcük dağarcığı örneğindeki verilerdir. Daha sonra birbirine bağımlı iki ortalamanın yer aldığı, yani tekrarlı ölçümler deseninin kullanıldığı bir durum incelenecektir.

Bağımsız Gruplar

Temel Kavram

Hatırlayacağınız gibi, üniversite öğrencilerinin ve yaşlı yetişkinlerin sözcük dağarcığı genişliklerinin değerlendirildiği bir çalışma yapılmıştı. Bu çalışmadaki duruma uygun çıkarımsal istatistik, **bağımsız gruplar için *t*-testidir.** Bu testi üniversite ve yaşlı yetişkin örneklemlerine ait, ortalama yüzde olarak ifade edilen çoktan seçmeli test performansları arasındaki farkı incelemek için kullanabiliriz. İstatistik yazılımları genellikle, elde edilen *t* değerine ait olasılığı çıktı ekranında vermektedir. *t* değerinin hesaplanması için gerekli olan formül, Ek A.2'dedir. APA *Yayın Kılavuzu* (2010), tam olasılığın rapor edilmesini önermektedir. Tam olasılık .001'den küçük olduğunda, (örneğin, $p = .0004$) istatistik yazılımları genellikle tam olasılığı .000 şeklinde rapor etmektedirler. (Yukarıda rapor edilen analizde de bu durum söz konusuydu.) Elbette tam olasılık .000 değil fakat .001'den küçük bir değerdir. Bağımsız gruplar *t*-testi için *serbestlik derecesi*, $n1 + n2 - 2$ şeklinde hesaplanır.

Buradan hareketle, incelemekte olduğumuz sözcük dağarcığı çalışmasına ait çıkarımsal istatistik sonucu şu şekilde özetlenebilir:

$$t(50) = 5.84, p < .001$$

Bölüm 11'de, iki ortalamanın karşılaştırılması amacıyla, etki büyüklüğü *d*'nin nasıl hesaplandığını göstermiştik. *Eğer SHAT kullanılıyorsa etki büyüklüğü ölçüsü her zaman rapor edilmelidir.* Bölüm 11'de, sözcük dağarcığı çalışması için *d* değerini 1.65 olarak hesapladığımızı hatırlıyor olabilirsiniz. Cohen *d*, aşağıdaki formül yardımıyla, bağımsız gruplar *t*-testi sonucundan hareketle de hesaplanabilir:

$$d = \frac{2t}{\sqrt{df}} \quad \text{(bk. Rosenthal ve Rosnow, 1991)}$$

Yani,

$$d = \frac{2(5.84)}{\sqrt{50}} = \frac{11.68}{7.07} = 1.65$$

Tekrarlı Ölçümler Desenleri

Temel Kavram

Buraya kadar, iki bağımsız denek grubunu içeren deneyleri ele aldık. Farkında olduğunuz gibi, aynı zamanda deneyler, deneklerin deneyin tüm koşullarına katılacağı biçimde ya da deneklerin, bağımlı değişken ile ilişkili olabilecek bir ölçüm üzerinde (örneğin, IQ, ağırlık) "eşleştirilmesi" biçiminde de yapılabilir. Bu tür deneyler, eşleştirilmiş gruplar (bk. Bölüm 6), denek içi desenler veya tekrarlı ölçümler deseni olarak adlandırılır (bk. Bölüm 7). Bununla birlikte, iki ortalamanın karşılaştırıldığı *t*-testi, tekrarlı ölçümler deseninde farklı bir şekil alır. Bu durumda *t*-testi, tipik olarak, doğrudan farklılık için *t* veya **tekrarlı ölçümler (denek içi) *t*-testi** olarak adlandırılır. Bilgisayar destekli bir analiz yapılırken, eğer denekler deneyin her iki aşamasında da yer alıyorsa, bağımsız denek gruplarının karşılaştırıldığı koşula göre verilerin daha farklı girildiğini fark edeceksiniz.

Tekrarlı ölçümler için t'nin payı, fark puanlarının ($\overline{D}$) ortalamasıdır ve cebirsel olarak örneklem ortalamaları arasındaki farka eşittir (yani $\overline{X}_1 - \overline{X}_2$). Payda ise, fark puanlarının tahminî standart hatasıdır (bk. Bölüm 11). İstatistiksel anlamlılık, elde edilen t ve t'nin, $N - 1$'e eşit olan sd'si için kritik değerlerinin karşılaştırılması ile belirlenir. Bu durumda N, katılımcı sayısına veya deneydeki puan çiftlerine karşılık gelir. Elde edilen t'yi, bağımsız gruplar desenlerinde elde edilen t'ye benzer şekilde yorumlayacaksınız.

Bölüm 11'de değinildiği gibi, etki büyüklüğünü eşleştirilmiş gruplar veya tekrarlı ölçümler deseninde değerlendirmek, bağımsız gruplar deseninde olduğundan biraz daha karmaşıktır (bu gibi durumlarda d'nin hesaplanması ile ilgili olarak daha fazla bilgi için bk., Cohen, 1988; Rosenthal ve Rosnow, 1991).

İSTATİSTİKSEL ANLAMLILIK VE BİLİMSEL YA DA PRATİK ANLAMLILIK

- İstatistiksel anlamlılığın bilimsel anlamlılıkla aynı şey olmadığı gerçeğini kabul etmeliyiz.
- Ayrıca istatistiksel anlamlılığın, pratik veya klinik anlamlılıkla da aynı şey olmadığını kabul etmeliyiz.

İstatistiksel anlamlılık testleri, araştırma bulgularının analizinde önemli birer araçtır. Bununla birlikte, istatistiksel olarak anlamlı bulguları doğru bir biçimde yorumlamaya dikkat etmeliyiz (bk. Kutu 12.3). Ayrıca istatistiksel olarak anlamlı bir bulguyu, bilimsel olarak anlamlı bir bulgu ile karıştırmamak konusunda da dikkatli olmalıyız. Bilim camiası için bir çalışmanın sonuçlarının önemli olup olmayacağı, çalışmanın altında yatan değişkenin niteliğine (basitçe, bazı değişkenlerin etkisi, diğerlerinden daha önemlidir), çalışmanın nasıl göründüğüne (uygun şekilde yürütülmemiş araştırmalarda, istatistiksel olarak anlamlı bulgular elde edilebilir) ve etki büyüklüğü gibi başka ölçütlere bağlıdır (örneğin, bakınız, Abelson, 1995).

Benzer şekilde, bir uygulama etkisinin pratik veya klinik anlamlılığı, istatistiksel anlamlılığın dışında bazı faktörlere bağlıdır. Bunların içerisinde çalışmanın dış geçerliği, etkinin büyüklüğü ve uygulamanın yürütülmesi ile ilgili çeşitli pratik meseleler (ekonomik olanlar da dâhil) yer alır. Yüksek etki gösteren ve istatistiksel olarak anlamlı olan bir sonuç bile, pratik veya klinik olarak anlamlılığın garantisini vermez. Çok yüksek bir etki büyüklüğü, laboratuvardan gerçek dünyaya genellemenin iyi bir biçimde yapılamadığı (yani düşük dış geçerlik) bir çalışmanın parçası olarak elde edilebilir; bu nedenle sonuçlar, uygulama yapan psikologlar için pek bir değer taşımayabilir. Dahası, gerçek dünyaya iyi bir şekilde genellenebilen, göreli olarak büyük bir uygulama etkisi çok maliyetli, uygulanması çok zor, çok tartışmalı veya var olan uygulamalar ile çok benzer etkiye sahip olduğu için hiçbir zaman uygulanamayabilir.

Ayrıca, düşük bir etki büyüklüğünün, eğer yeterli güç varsa, istatistiksel olarak anlamlı olması da olasıdır. Küçük etki büyüklüklerinin, laboratuvar dışında pratik bir önemi olmayabilir. Bölüm 6'da açıkladığımız gibi, dış geçerlik görgül bir meseledir. Bir bulgunun pratik olarak anlamlı olup olmayacağını görebilmek için çalışmayı, uygulamanın yapılacağı koşullara benzer koşullarda yapmak önemlidir. Bununla birlikte, etki büyüklüğü küçükse, böyle bir görgül test gerçekleştirmemiz pek olası değildir (önemli istisnalar için bk. Rosenthal, 1990).

KUTU 12.3

BİR SONUÇ İSTATİSTİKSEL OLARAK ANLAMLI OLDUĞUNDA ($p < .05$) NE SÖYLEMEMELİYİZ?

- Ortalamalar arasındaki gerçek farka ilişkin tam olasılığı belirleyemeyiz. Örneğin, .95 olasılıkla ortalamalar arasında gözlenen farkın, evrenlerin ortalamaları arasındaki gerçek farkı yansıttığını söylemek doğru değildir.

SHAT'nin sonucu, gözlenen farkın, sıfır hipotezini doğru varsayarak, şans eseri oluşma olasılığını (eldeki verilerden hareketle) gösterir. SHAT bize gerçek dünyadaki olasılıklarla ilgili bir şey söylemez (örneğin, Mulaik ve ark., 1997). Eğer sonuçlar seçtiğimiz alfa düzeyinden (örneğin, .05) daha düşük bir olasılıkla ortaya çıkıyorsa, elde edilen bulgunun, bu durumda şans eseri oluşmadığı sonucuna varabiliriz.

- İstatistiksel olarak anlamlı bir sonuç, araştırma hipotezinin doğru olduğunu göstermez (Örneğin sözcük dağarcığı çalışmasındaki veri, yaşlı yetişkinlerin genç yetişkinlere göre daha geniş bir sözcük dağarcığı olduğunu kanıtlamaz.).

SHAT (güven aralıklarında olduğu gibi) bir araştırma hipotezinin doğru olduğunu kanıtlayamaz. İstatistiksel olarak anlamlı bir sonuç (anlaşılabilir şekilde), bazen hipotezle ilgili olarak "desteklemektedir" ya da "kanıt sunmaktadır" şeklinde yorumlanır ama söz konusu sonuç tek başına araştırma hipotezinin doğru olduğunu kanıtlayamaz. Bunun neden böyle olduğuyla ilgili birkaç önemli sebep vardır. Öncelikle SHAT, bir tür olasılık oyunudur; SHAT, değeri hiçbir zaman 1.00 olmayan olasılıklar şeklinde yanıtlar verir (örneğin, p .05'ten büyük ya da küçük). Her zaman hata yapma olasılığı vardır. Eğer bir "kanıt" söz konusu ise, bu sadece "ikinci dereceden (dolaylı)" bir kanıttır. Daha önce gördüğümüz gibi, araştırma hipotezi, sadece sıfır hipotezinin doğru olduğu varsayımı altında, eldeki veriye atıfta bulunularak dolaylı şekilde test edilebilir. Eğer sonuçlarımızın şans eseri ortaya çıkmış olma olasılığı çok düşükse (sıfır hipotezinin doğru olduğunu varsayarak), sıfır hipotezinin doğru olmadığı yönünde muhakeme yapabiliriz; bununla birlikte bu durum araştırma hipotezimizin doğru olduğu anlamına gelmez. Schmidt ve Hunter'ın (1997, s. 5) hatırlattığı gibi, SHAT uygulayan araştırmacılar "asıl ilgilenilen bilimsel hipoteze odaklanmaz." İkinci olarak, bağımsız değişkenin etkisine dair kanıt, ancak etkiyi doğuran yöntem kadar güçlü olabilir. Gözlenen etkinin dışında, başka bir faktörün sorumlu olması olasıdır (Örneğin, sözcük dağarcığı çalışmasındaki yaşlı yetişkinlerin çapraz bulmaca ustaları arasından seçildiğini, fakat üniversite öğrencilerinin bu şekilde seçilmediğini düşünün.). Daha önce belirttiğimiz gibi, kötü tasarlanmış bir deney, kolayca yüksek bir etki büyüklüğü sağlayabilir. Araştırma hipotezi için elde edilen bir kanıt, bağımlı değişken üzerinde elde edilen etki büyüklüğü kadar, çalışmanın yöntemi incelenerek de aranmalıdır. Ne SHAT ve güven aralıkları ne de etki büyüklükleri, çalışmanın yönteminin sağlamlığına dair bize bir şey söylemez.

İKİ ORTALAMANIN KARŞILAŞTIRILMASI İLE İLGİLİ ÖNERİLER

İki ortalama arasındaki farkı inceleyen bir çalışmadan elde edilen verileri değerlendirmekle ilgili şu önerileri getirebiliriz. Öncelikle, veri analizinin en son amacını aklınızdan çıkarmayın: Gözlemlerinize dayanarak, davranışa ilişkin bir iddianın gerekçesini oluşturmak. En iyi gerekçeyi oluşturmak amacıyla, veri analizi için farklı alternatifleri incelemeye gereksinim duyacaksınız. Davranışa ilişkin bir iddiaya kanıt elde etmek için sadece bir yol olduğu düşüncesi bir tuzaktır; bu tuzağa düşmeyin. Eğer seçme şansınız varsa (neredeyse her zaman vardır), APA'nın İstatistiksel Çıkarım Çalışma Kolu'nun (Wilkinson ve ark., 1999) önerdiği gibi, kullanılabilecek en basit analizi kullanın. İkinci olarak,

SHAT'yi kullanırken, bu analizin sınırlılıklarını ve SHAT'nin sonucunun, sizin neyi söylemenize izin verdiğini anladığınızdan emin olun. SHAT kullanırken etki büyüklüğü ölçüsünü ve aynı zamanda da gücünü, özellikle sonuçlar anlamlı olmadığında, daima rapor ediniz. Her ne kadar etki büyüklüğüne ilişkin bilginin garanti edilemediği durumlar olabilirse de –örneğin, sadece yönün kuramsal bir tahminini yapmakta olduğu gibi (örneğin, Chow, 1988); buna benzer durumlara görece ender rastlanmaktadır. Birçok araştırmada ve hemen hemen tüm uygulamalı araştırmalarda etki büyüklükleri SHAT'nin önemli, hatta gerekli bir tamamlayıcısıdır. Son olarak, araştırmacılar sadece SHAT'ye güvenme "alışkanlığından kurtulmalı" ve p değerlerine ek olarak, hatta onların yerine, etki büyüklükleri için güven aralıklarını rapor etmelidirler. APA *Yayın Kılavuzu* (2010, s. 33), güven aralıklarının kullanılmasını ısrarla önermektedir.

İKİ ORTALAMAYI KARŞILAŞTIRIRKEN SONUÇLARIN RAPOR EDİLMESİ

Şimdi, veri analizinin her üç aşamasından elde edilen bilgiyi, güven aralıklarının kullanılmasıyla elde edilen tamamlayıcı kanıtı ve SHAT'yi, ayrıca APA *Yayın Kılavuzu'nun* (2010) sonuçların rapor edilmesiyle ilgili önerilerini (özellikle kılavuzdaki ss. 32–35'e bakınız) dikkate alan, bulgulara dair bir ifadeyi hazırlayabilecek durumdayız. SHAT ve güven aralığını (sonuçlar bölümünde *GA* olarak kısaltılmaktadır) kullanarak sonuçların rapor edilmesiyle ilgili ek bilgi Bölüm 13'tedir.

Sözcük Dağarcığı Çalışmasına Ait Sonuçların Rapor Edilmesi Sonuçları şu şekilde rapor edebiliriz:

> Üniversite öğrencileri için, çoktan seçmeli sözcük dağarcığı testinden elde edilen ortalama performans 45.58 (S = 10.46), yaşlı grubunun ortalaması ise 64.04'tür (S = 12.27). Bu fark istatistiksel olarak anlamlıdır; $t(50) = 5.84$, $p < .001$, $d = 1.65$, %95 *GA* [12.11, 24.28]. Araştırmada, yaşlı katılımcıların daha geniş bir sözcük dağarcığına sahip olduğu bulunmuştur.

Yorum Ortalamalar ve standart sapmalar şeklinde sunulan betimsel istatistikler, deneyde, bağımsız değişkenin (yaş) bir fonksiyonu olarak "nelerin gerçekleşmiş olduğunu" özetlemektedir. Tam olasılık .001'den küçük olduğu için, sonuçlar $p < .001$ şeklinde rapor edilmektedir ama tam olasılıkların değeri .001 veya üzeri olduğu zaman, olasılıkların rapor edilmesi gerektiğine dikkat edin. Tam olasılık, çalışmanın tam bir tekrarının yapılabilme olasılığı hakkında bilgi verir (Posavac, 2002). Başka bir deyişle, daha yüksek bir p değerinin elde edildiği duruma göre, sonuçların "daha güvenilir" olduğunu biliriz. Sadece güven aralıkları rapor edildiği zaman, bu bilgi kaybedilir. "Bu çalışmadaki yaşlı yetişkinler..." şeklinde başlayan cümle, istatistiksel analizin ne anlama geldiğini kelimelerle ifade eder. Analiz sonuçlarının ne anlama geldiğini okuyucunuza doğrudan anlatmak her zaman önemlidir. Bu durumun önemi, araştırmada yapılan ve rapor edilen analizlerin karmaşıklığı ve miktarı arttıkça, daha da artar. APA *Yayın Kılavuzu'nun* önerdiği gibi, etki büyüklüğü de (yani d) rapor edilir. Bu bilgi, meta-analiz yapan araştırmacılar ile birbirine benzer değişkenleri kullanan çalışmaların sonuçlarını karşılaştırmak isteyenler için değerlidir. Öte yandan güven aralıkları, Cohen d'de olduğu gibi tek bir değer değil, asıl ortalama farkları çerçevesinde, olası etki büyüklüklerine dair bir aralık sunmaktadır. Sıfır,

değeri aralığının içerisinde yer almadığı için, biliyoruz ki sonuç istatistiksel olarak .05 düzeyinde anlamlıdır (bk. Bölüm 11). Bununla birlikte, APA *Kılavuzu*'nun vurguladığı gibi güven aralıkları, SHAT'nin tek başına vermediği tahminin dakikliğine ve etkinin yer aldığı bölgeye ilişkin bilgileri verir. Daha dar güven aralıklarının daha kesin tahmin yapabilmemiz anlamına geldiğini Bölüm 11'den hatırlayınız.

Güç Analizi Etki büyüklüğünün değerini bildiğimiz zaman, güç tablolarını kullanarak analizin istatistiksel gücünü belirleyebiliriz. Hatırlayacağınız gibi güç, istatistiksel olarak anlamlı bir etkinin elde edilebilme olasılığıdır. Genç ve yaşlı yetişkinlerin karşılaştırıldığı sözcük dağarcığı çalışmasında, Cohen'in (1988) kriterlerine göre orta düzeyde olan .50 büyüklüğünde bir etki elde edildiğini varsayın. .05 düzeyindeki bir alfa söz konusu olduğunda, .50 büyüklüğünde bir etkiyi "görmek" için, ortalamalar arası farka ilişkin bir testte gerekli olan katılımcı sayısını belirlemek amacıyla Cohen'in oluşturduğu güç tablolarını kullanabiliriz. Bunun sonucunda, iki gruplu bir çalışma için örneklem büyüklüğünün (her bir grup için), .80'lik bir güce ulaşabilmek amacıyla, 64 olması gerektiği görülmektedir (iki uçlu test için). Orta düzeyde bir etki amaçlandığında, 10 denemenin 8'inde istatistiksel olarak anlamlı bir sonuç elde edebilmek için toplamda 128 (64 × 2) katılımcıya ihtiyacımız vardır. Araştırmacılar orta düzeyde bir etkiyi arıyor olsaydı, araştırmacıların sözcük dağarcığı çalışması, olduğundan daha az bir güce sahip olacaktı. Fakat görüldüğü gibi, yüksek bir etki gücü beklendiğinde, 26 kişilik bir örneklem, .80'lik bir güç elde etmek için uygun olacaktır.

Sonuç istatistiksel olarak anlamlı değilse, güç tahmini rapor edilmelidir. Örneğin, bağımsız gruplar deseni kullanıldığında sonuç $t(28) = 1.96$, $p > .05$ şeklinde ve etki büyüklüğü .50 ise, çalışmanın gücünü sonuç elde edildikten sonra belirleyebiliriz. Çalışmadaki grupların eşit büyüklükte olduğunu varsayarsak her grupta 15 denek olduğunu biliyoruz ($sd = n_1 + n_2 - 2$, ya da $28 = 15 + 15 - 2$). Güç analizi, bu çalışmanın gücünün .26 olduğu sonucunu verecektir. İstatistiksel olarak anlamlı bir sonuç, bu büyüklükteki bir örneklem ve orta düzeyde bir etki (.50) için, yaklaşık olarak dört denemenin sadece birisinde bulunabilecektir.

İLERİ ALIŞTIRMA
SIFIR HİPOTEZİNİN TEST EDİLMESİNE İLİŞKİN BİR TEST (KONUYU NE KADAR İYİ ANLADIĞINIZA İLİŞKİN BİR TEST)

Anlaşılmış olacağı gibi, SHAT'yi anlamak, uygulamak, ve SHAT'nin sonuçlarını yorumlamak kolay bir iş değildir. Deneyimli araştırmacılar bile zaman zaman hata yapabilmektedir. Hatalardan kaçınmak için size yardımcı olmak amacıyla, SHAT hakkında şimdiye kadar verilen bilgilerden hareketle, bir doğru-yanlış testi hazırladık.

Deney ve kontrol gruplarında yer alan katılımcıların performansını değerlendirmek amacıyla bağımsız gruplar deseninin kullanıldığını varsayın. Her koşulda 12 katılımcı yer almaktadır ve .05 düzeyinde alfa kullanılarak yapılan SHAT'nin sonucu, $t(22) = 4.52$, $p = .006$'dır. Bu sonuçtan hareketle, araştırmacının varacağı aşağıdaki sonuçların hangileri doğru, hangileri yanlıştır?

1. Sıfır hipotezi reddedilmelidir.
2. Araştırma hipotezinin doğru olduğu gösterilmiştir.
3. Sonuçlar bilimsel açıdan önemlidir.
4. Sıfır hipotezinin doğru olma olasılığı sadece .006'dır.
5. Tam olasılık .02 olsaydı, çalışma tekrarlandığında .05 düzeyinde istatistiksel olarak anlamlı bir sonuç bulma olasılığı da yüksek olacaktı.

Bu durumda, elde edilen sonuç temel alınarak, pratik veya kuramsal bir karara varılıp varılamayacağına veya "daha fazla araştırmanın yapılmasına gereksinim olup olmadığına" araştırmacıların karar germesi gerekir.

İKİDEN FAZLA KOŞUL İÇİN VERİ ANALİZİ

Şimdiye kadar, iki koşulu olan bir deney bağlamında, başka bir deyişle, iki düzeyi olan bir bağımsız değişken çerçevesinde veri analizinin aşamalarını tartıştık. Peki, ikiden fazla düzeyimiz (koşulumuz) olduğunda, veya psikolojide sık karşılaşıldığı gibi, ikiden fazla bağımsız değişkenimiz olduğunda durum ne olur? Bu gibi durumlarda, psikoloji deneylerinden elde edilen sonuçları analiz etmek için en sık kullanılan istatistiksel işlem varyans analizidir (ANOVA).

Dört özel araştırma durumunda, sıfır hipotezini test etmek için ANOVA'nın nasıl kullanıldığını göstereceğiz: Bağımsız gruplar deseni için tek faktörlü analiz, tekrarlı ölçümler deseni için tek faktörlü analiz, bağımsız gruplar deseni için iki faktörlü analiz ve karma desenler için iki faktörlü analiz. Devam etmeden önce, Bölüm 6, 7 ve 8'de yer alan ve bu araştırma desenlerini açıklayan bilgileri gözden geçirmenizi öneriyoruz.

TEK FAKTÖRLÜ BAĞIMSIZ GRUPLAR DESENİ İÇİN ANOVA

- Varyans analizi (ANOVA), bir bağımsız değişkenin, bağımlı değişken üzerinde istatistiksel olarak anlamlı etkisinin olup olmadığını incelemek için kullanılan bir çıkarımsal istatistik testidir.
- Varyans analizinin altında yatan mantık, hata değişkenliği ve sistematik değişkenliğe ilişkin kaynakların belirlenmesine dayanır.
- *F*-testi, veriye ait gruplar arası değişkenlik ile grup içi değişkenlik arasındaki oranı veren bir istatistiktir.
- Başlangıçta yapılan genel *F*-testi analizinin sonuçları, varyans analizi özet tablosunda sunulmaktadır.
- Varyans analizi, bağımsız değişkenin istatistiksel olarak anlamlı bir etkisinin olup olmadığına karar vermek için kullanılabilirse de araştırmacılar deneyin sonucunu yorumlayabilmek için betimsel istatistikleri incelerler.
- Bağımsız gruplar deseni için olan etki büyüklüğü ölçüleri arasında eta kare (η^2) ve Cohen *f* yer alır.
- Bağımsız gruplar deseni için güç analizi, istatistiksel olarak anlamlı bir etki bulma olasılığını belirlemek amacıyla araştırmaya başlamadan önce yapılmalıdır ve SHAT kullanılarak anlamlı olmayan sonuçlar elde edildiğinde güç rapor edilmelidir.
- Ortalamaların ikili olarak karşılaştırılması, istatistiksel olarak anlamlı olan genel *F*-testine katkıda bulunan sistematik değişkenliğin spesifik kaynaklarının belirlenmesi için kullanılabilir.

Temel Kavramlar

Genel Bakış İstatistiksel çıkarım, deney sonucunun istatistiksel olarak anlamlı olup olmadığını belirleyen testlerin kullanımını gerektirir. Psikoloji deneylerinin analizinde en sık kullanılan istatisiksel çıkarım testi **ANOVA**'dır. Adından da anlaşılacağı gibi varyans analizi, bir deneydeki değişkenliğin farklı kaynaklarının analizini temel alır. Bu bölümde, bağımsız grupların yer aldığı ve bir bağımsız değişkeni olan deneylerin, bilinen adıyla, **tek faktörlü bağımsız gruplar deseninin** analizi için varyans analizinin nasıl kullanıldığını kısaca ele alacağız. ANOVA hem seçkisiz gruplar hem de doğal gruplar deseni için kullanılıyor olsa da, ANOVA'nın temelindeki varsayımlar katı bir biçimde sadece seçkisiz gruplar deseni için geçerlidir.

Herhangi bir seçkisiz gruplar deseninde değişkenliğe ait iki kaynak vardır. Öncelikle, gruplara seçkisiz olarak atanan denekler arasındaki bireysel farklılıklar nedeniyle her bir grubun kendi içinde değişkenlik beklenebilir. Bireysel farklılıklardan kaynaklanan değişkenlik ortadan kaldırılamaz ama seçkisiz atama yapıldığı zaman bu değişkenliğin gruplar arasında dengelendiği kabul edilir. Uygun şekilde yapılmış olan bir deneyde her bir grup içindeki denekler arasındaki farklılıklar, hata değişkenliğinin tek kaynağı olmalıdır. Her gruptaki katılımcılara yönergeler aynı şekilde verilmeli ve katılımcıların atandıkları bağımsız değişken düzeyi, grubun her üyesine aynı şekilde uygulanmalıdır (bk. Bölüm 6).

Seçkisiz gruplar deseninde değişkenliğin ikinci kaynağı, gruplar arasındaki değişkenliktir. Sıfır hipotezi doğru ise (gruplar arasında fark yok ise), grup ortalamaları arasında gözlenen herhangi bir fark, hata değişkenliğine bağlanabilir (örneğin, gruplarda yer alan katılımcıların farklı özelliklere sahip olması gibi). Bununla birlikte, daha önce de gördüğümüz gibi, örneklem ortalamalarının tamamen aynı olmasını beklemeyiz. Örnekleme hatasından kaynaklanan dalgalanmalar ortalamaların bir miktar farklılaşmasına neden olur – bu da hata değişkenliğidir. Bu nedenle, farklı grupların ortalamaları arasındaki değişkenlik, sıfır hipotezinin doğru olduğu varsayıldığında, bir deneydeki hata değişkenliğine ilişkin ikinci bir tahmin verir. Eğer sıfır hipotezi doğruysa, *gruplar arasındaki* hata değişkenliği ile ilgili bu tahmin, grup içindeki hata değişkenliğine ilişkin tahminle aynı olmalıdır. Bu nedenle seçkisiz gruplar deseni, hata değişkenliğine ait iki bağımsız tahmin sunar; bunların biri gruplar içinde, diğeri de gruplar arasındadır.

Şimdi de sıfır hipotezinin yanlış olduğunu varsayın. Başka bir deyişle, deneyinizdeki bağımsız değişkenin bir etkisinin olduğunu varsayın. Eğer bağımsız değişkenin bir etkisi varsa, farklı gruplara ait ortalamaların farklılaşması gerekir. Davranış üzerinde etkisi olan bir bağımsız değişken, deneydeki farklı grupların ortalamaları arasında sistematik farklar doğurmalıdır. Başka bir deyişle bağımsız değişken, deneydeki gruplar arasındaki değişkenliğe dair bir kaynak sunmalıdır –grupların farklılaşmasına neden olmalıdır. Bu sistematik değişkenlik, grup ortalamaları arasında, hata değişkenliği nedeniyle zaten var olan farklara eklenecektir. Yani gruplar arası değişkenlik artacaktır.

***F*-Testi** Şimdi, bağımsız değişkenimizden kaynaklanan değişkenliğin, sadece hata değişkenliği nedeniyle beklenebilecek değişkenlikten daha büyük olup olmadığını öne sürmemizi sağlayacak bir istatistik kullanabilecek durumdayız.

Temel Kavram

Bu istatistiğe *F* denir; bu adı, testi geliştiren istatistikçi Ronald Fisher'den almıştır. ***F*-testinin** kavramsal tanımlaması aşağıdaki şekildedir:

$$F = \frac{\text{Gruplar arası değişkenlik}}{\text{Grup içi değişkenlik}} = \frac{\text{Hata değişkenliği + Sistematik değişkenlik}}{\text{Hata değişkenliği}}$$

Sıfır hipotezi doğru ise gruplar arasında sistematik bir değişkenlik yok demektir (bağımsız değişkenin bir etkisi yoktur) ve *F*-testinin beklenen sonucu 1.00'dir (hata değişkenliğinin hata değişkenliğine bölümü 1.00 olacağı için). Bununla birlikte, sistematik değişkenliğin miktarı arttıkça, *F*-testinden beklenen değer de 1.00'ın üzerine çıkar.

Eğer bağımsız değişkenin ürettiği sistematik değişkenliği ayırabilseydik, deneyleri analiz etmek daha kolay olurdu. Ne yazık ki gruplar arasındaki sistematik değişkenlik, hata değişkenliği ile birlikte bir "paket" hâlinde ortaya çıkmaktadır. Sonuç olarak, *F*-testinin sonucu basitçe, gruplar arasındaki hata tahminimizin, gruplar içindeki hata tahminimizden daha büyük olması nedeniyle, bazen 1.00'dan büyük olabilmektedir (yani, iki tahmin aynı olmalıdır fakat şansa bağlı faktörler nedeniyle farklılaşabilmektedir). Peki, *F* istatistiğinin bağımsız değişken nedeniyle oluşan gerçek sistematik değişkenliği yansıttığından göreli olarak emin olmadan önce, bu istatistiğin 1.00'dan ne kadar büyük olması gerekir? İstatistiksel anlamlılık üzerine önceden yaptığımız tartışma bu soruya cevap vermektedir. *F* değerinin istatistiksel olarak anlamlı olabilmesi için, sıfır hipotezi doğru olduğunda bu değerin ortaya çıkma olasılığının, seçtiğimiz anlamlılık değerinden, genellikle .05'ten küçük olmaya yetecek kadar büyük olması gerekir.

Şimdi, belirli bir deneyi analiz etmek için SHAT'nin ilkelerini ve ANOVA işlemlerini uygulamaya hazırız.

Tek Faktörlü Bağımsız Gruplar Deseninin Analizi *F*-testi gibi bir çıkarımsal istatistiksel testi yapabilmek için ilk adım, analizin yanıtlamayı amaçladığı araştırma sorusunu ifade etmektir. Bu soru genellikle şöyle sorulur: "Bağımsız değişkenin performans üzerinde genel bir etkisi var mı?". Araştırma sorusu netleştiğinde ikinci adım, analiz için bir sıfır hipotezi ortaya atmaktır. Örnek olarak tartışacağımız deney, çeşitli bellek eğitimlerinin, bilginin bellekte tutulması üzerindeki etkisini incelemektedir. Bu bağımsız değişkenin dört düzeyi (koşulu) bulunmaktadır ve dolayısıyla dört katılımcı grubu yer almaktadır. Her örneklem, yani grup, bir evreni temsil etmektedir. Deneyin ilk genel analizine, **genel *F*-testi** denir. Bu tür genel testler için kurulan sıfır hipotezi, tüm evrenlerin ortalamalarının eşit olduğunu öne sürer. Sıfır hipotezinin, bağımsız değişkenin bir etkisinin olmadığını öne sürdüğünü anımsayın. Bir sıfır hipotezi (H_0), daima evren özelliklerine göre ifade edilir. Bu evren özellikleri Yunan harfleri ile gösterilir ve evren ortalaması da μ ("mü") ile sembolize edilir. Bağımsız değişkenin düzeylerini göstermek amacıyla her ortalama için bir alt simge kullanabiliriz. Bu durumda sıfır hipotezimiz aşağıdaki gibi olur:

Temel Kavram

$$H_0: \mu_1 = \mu_2 = \mu_3 = \mu_4$$

Sıfır hipotezinin alternatifi, evrenlerin bir veya daha fazla ortalamasının eşit olmadığını öne sürer. Başka bir deyişle, alternatif hipotez (H_1), H_0'ın yanlış

TABLE 12.2 BİR BELLEK DENEYİNDE ANIMSANAN SÖZCÜK SAYISI

	Yönerge (A)						
Denek	Kontrol (a_1)	Denek	Hikaye (a_2)	Denek	İmgeleme (a_3)	Denek	Kafiye (a_4)
1	12	6	15	11	16	16	14
2	10	7	14	12	16	17	14
3	9	8	13	13	13	18	15
4	11	9	12	14	12	19	12
5	8	10	12	15	15	20	12
Ortalama	10.0		13.2		14.4		13.4
Standart sapma	1.6		1.3		1.8		1.3
Aralık	8–12		12–15		12–16		12–15

olduğunu, yani fark olduğunu öne sürer. Bu durumda alternatif hipotez, aşağıdaki gibi şekillenir:

$$H_1: \text{DEĞİL } H_0$$

Eğer bellek eğitiminin bellekte tutma üzerinde bir etkisi varsa (yani bağımsız değişken sistematik değişkenlik doğuruyorsa), bu durumda sıfır hipotezini reddetmek isteriz.

Tablo 12.2'deki veriler, bellek eğitimi tekniklerini araştıran bir deneyde 20 soruluk bir bellek testinde doğru anımsanabilen sözcüklerin sayısıdır. Deneyde her bir gruba seçkisiz olarak beş katılımcı atanmıştır (gruplar, sözcükleri öğrenmek için katılımcıların kullandıkları çalışma yöntemine göre belirlenmiştir). Kontrol yöntemine herhangi bir yönerge verilmemiştir ama üç deney grubuna, sözcükleri öğrenebilmeleri için üç farklı yönerge verilmiştir: Öğrenme listesindeki sözcükleri kullanarak bir öykü uydurmaları (öykü yöntemi); görsel imgeleme kullanmaları (imgeleme yöntemi) veya kafiye kullanmaları (kafiye yöntemi). Bu deneyde değişimlenen bağımsız değişken, "yönerge türü"dür ve "A" harfi ile gösterilebilir. Bağımsız değişkenin düzeyleri dört grup için sırasıyla, a_1, a_2, a_3 ve a_4 simgeleri kullanılarak birbirinden ayrılabilir. Her gruptaki katılımcı sayısı n ile gösterilmektedir; buradaki durumda, $n = 5$. Deneydeki toplam katılımcı sayısı N ile gösterilmektedir; burada $N = 20$'dir.

Herhangi bir deneyin analizindeki önemli bir adım, Tablo 12.2'deki gibi bir veri matrisi hazırlamaktır. Tabloda, her dört gruptaki her bir birey için doğru tepki sayısı listelenmiş ve her katılımcı farklı bir denek numarası ile belirtilmiştir. Bir deneyin sonucunu anlayabilmek için, ANOVA sonucunu incelemeye başlamadan önce verileri özetlemek çok önemlidir. Veri matrisinin altında her grup için ortalama, aralık (en düşük ve en yüksek puanlar) ve standart sapma verilmiştir.

Herhangi bir çıkarımsal testin "anlamlılığını" incelemeden önce, özet istatistiklerin size ne söylediğine dair bir izlenim edinmeye çalışın. Bağımsız değişkenin görünürde bir etkisinin olup olmadığını inceleyin; yani ortalamalar arasında

TABLO 12.3 BELLEK DENEYİNE AİT VARYANS ANALİZİ ÖZET TABLOSU

Kaynak	Kareler Toplamı	*sd*	Ortalama Kare	*F*-Oranı	*p*
Grup	54.55	3	18.18	7.80	0.002
Hata	37.20	16	2.33		
Toplam	91.75	19			

belirgin bir değişkenliğin var olup olmadığına bakın. Aralıkları ve standart sapmaları inceleyerek, her gruptaki değişkenlik hakkında bir fikir edinin (Puanların örneklem ortalaması etrafında gösterdikleri değişkenlik azaldıkça, var olan bir etkiyi görebilme olasılığının arttığını anımsayın.). Aralık, yani en düşük ve en yüksek değerler arasındaki fark, taban ve tavan etkilerinin belirlenmesinde kullanışlıdır. Peki, gruplar arasındaki değişkenlik benzer midir? Denek içi değişkenlikteki büyük farklılaşmalar, ANOVA kullanırken yorumlama sorunlarına yol açtığı için değişkenliğin göreli olarak homojen olmasını isteriz.

Özet istatistikler üzerindeki incelememiz ortalamalar arasında sistematik bir değişkenlik olduğunu göstermektedir; en büyük fark Kontrol Grubu (10.0) ve İmgeleme Grubu (14.4) arasındadır. Tüm deney koşullarına ait ortalamalar, kontrol koşulu ortalamasından büyüktür. Tüm gruplar için aralığın benzer olduğuna dikkat edin; standart sapmalar da oldukça benzerdir. Bu durum, gruplar arasında değişkenlik benzerliğinin (homojenliğinin) var olduğunu göstermektedir (Birçok bilgisayar programı, ANOVA sonucuyla birlikte "varyans homojenliği" testini de vermektedir.). Buna ek olarak, her gruptaki en yüksek puanların incelenmesi, bu veri kümesinde tavan etkilerinin bir sorun yaratmadığını göstermektedir (çünkü olası en yüksek değer 20'dir).

Varyans analizinde bir sonraki adım, *F*-testinin pay ve paydasını oluşturan değişkenliğe ait tahminleri elde etmek için gerekli hesaplamaları yapmaktır. *F*-testine ilişkin hesaplamalar bilgisayarla kolayca yapılmaktadır. Bu nedenle daha çok hesaplamaların yorumlanmasına odaklanacağız. Varyans analizi sonuçları, *Varyans Analizi Özet Tablosunda* sunulmuştur (bk. Tablo 12.3).

ANOVA Özet Tablosunun Yorumlanması Tablo 12.3'te, bellek eğitiminin etkisini incelemek amacıyla kullanılan bağımsız gruplar deseni için Genel *F*-testine ait özet tablo yer almaktadır. $n = 5$ büyüklüğünde dört grup olduğunu (dolayısıyla $N = 20$) hatırlayın. ANOVA Özet Tablosunun ne içerdiğini biliyor olmanız çok önemlidir. Bu nedenle deneye ilişkin *F*-testinin sonucuna bakmadan önce, özet tablosunun bileşenlerini inceleyeceğiz.

Özet tablosunun sol sütunu, önceden sözü edilen iki değişkenlik kaynağını göstermektedir. Buradaki örnekte eğitim grubu bağımsız değişkeni ("Grup") gruplar arası değişkenliğin kaynağıdır ve grup içi farklılıklar da hata değişkenliğinin bir tahminidir. Deneydeki toplam değişkenlik ise gruplar arası ve grup içi değişkenliğin toplamıdır. Üçüncü sütun, serbestlik derecesini (*sd*) göstermektedir. Genelde serbestlik derecesi, istatistiksel bir kavram olarak, ilgilenilen girdi sayısı – 1 biçiminde tanımlanır. Eğitim bağımsız değişkeninin 4 düzeyi olduğu için gruplar arası *sd*, 3'tür. Her grupta 5 katılımcı yer almaktadır; bu nedenle 4 grubun her biri içinde *sd*, 4'tür ($n - 1$). Dört grup da eşit büyüklükte olduğu için, her bir grup içindeki *sd*'yi grup sayısı ile çarparak (4 x 4), 16 *sd*

olarak belirleyebiliriz. Toplam *sd*, denek sayısı – 1'dir (N – 1), veya gruplar arası *sd* ve grup içi *sd*'nin toplamıdır (3 + 16 = 19).

Kareler toplamı (*KT*) ve ortalama kare (*OK*), *F* istatistiğinin elde edilmesindeki sayısal adımlardır. Gruplar arası *OK* (satır 1), sistematik değişkenlik ile hata değişkenliğinin toplamına ait bir tahmindir ve gruplar arası *KT*'nin gruplar arası *sd*'ye bölünmesi (54.55/3 = 18.18) ile hesaplanır. Grup içi *OK* (satır 2), sadece hata değişkenliğinin bir tahminidir ve grup içi *KT*'nin grup içi *sd*'ye bölünmesi (37.20/16 = 2.33) ile hesaplanmaktadır. *F*-testi ise, gruplar arası *OK*'nin, grup içi *OK*'ye bölünmesi (18.18/2.33 = 7.80) ile hesaplanmaktadır.

Şimdi, bellek eğitimi deneyindeki sonuca ilişkin istatistiksel anlamlılığı test etmek üzere, özet tablosunda yer alan bilgiyi kullanmaya hazırız. Sıfır hipotezinin doğru olduğu varsayıldığında (yani bağımsız değişkenin bir etkisi olmadığında), sistematik değişkenliğin ve hata değişkenliğinin toplamına ilişkin tahminin (*F*-testinin payı), tek başına hata değişkenliği tahminine (*F*-testinin paydası) yaklaşık olarak eşit olması gerektiğini bildiğiniz için sonucu zaten tahmin ediyor olabilirsiniz. Burada gördüğümüz gibi, sistematik değişkenlik ve hata değişkenliğinin toplamına ait tahmin (18.18), tek başına hata değişkenliği tahmininden (2.33) epeyce büyüktür.

Bu analizde elde edilen *F* değeri (7.80), özet tablonun sondan bir önceki sütununda görülmektedir. Eğer sıfır hipotezi doğru olsaydı, 7.80 büyüklüğündeki bir *F* değerini elde etme olasılığı, özet tablosunun son sütununda gösterilmektedir (0.002). Elde edilen .002 olasılığı, anlamlılık düzeyinden (α = .05) küçüktür ve bu durumda sıfır hipotezini reddedip, bellek eğitiminin genel etkisinin istatistiksel olarak anlamlı olduğu sonucuna varırız. Araştırma raporunuzda, ANOVA kullanarak yaptığınız SHAT'nin sonucu şu şekilde özetlenecektir:

$$F(3, 16) = 7.80, p = .002$$

F istatistiği, serbestlik derecesi ile belirlenir. Buradaki örnekte gruplar arası *sd* 3, grup içi *sd* 16'dır (yani 3, 16). Sonuçların tekrar edilebilirliğinin olasılığını verdiği için tam olasılığın (yani .002) rapor edildiğine dikkat edin.

Genel bir sıfır hipotezini test eden varyans analizinde istatistiksel olarak anlamlı bir sonuç bulduğumuzda ne öğrenmiş olduk? Bir açıdan, çok önemli bir şeyi öğrendik. Şu anda, bağımsız değişkeni değişimlemenin performansta (yani katılımcıların, hatırlanacak sözcüklere ait belleğinde) bir değişiklik yarattığını söyleyebilecek durumdayız. Bir başka açıdan ise, yalnızca sonucumuzun istatistiksel açıdan anlamlı olduğunu biliyor olmak, bize, bağımsız değişkene ait etkinin niteliği ile ilgili çok az şey söyler. Betimsel istatistikler (bizim örneğimizde, Tablo 12.2'de rapor edildiği gibi, anımsanan sözcük sayısı ortalaması) etkinin niteliğini betimlememizi sağlar. Deneyimizde, bağımsız değişkenin bir fonksiyonu olarak nelerin gerçekleşmiş olduğunu, sadece grup ortalamalarına ilişkin örüntüyü inceleyerek öğrenmeye başlayabiliriz. *Hiçbir zaman, istatistiksel olarak anlamlı olan bir sonucu, ilgili betimsel istatistiklere atıfta bulunmadan yorumlamaya çalışmayın.*

Her ne kadar genel *F*-testinin istatistiksel olarak anlamlı olduğunu biliyor olsak da, bağımsız ve bağımlı değişkenler arasındaki ilişkinin derecesi hakkında bir şey bilmiyoruz ve bu yüzden, bağımsız değişkenimizin etki

büyüklüğünü hesaplamayı göz önünde bulundurmalıyız. Sadece genel teste dayanarak, hangi grup ortalamalarının anlamlı düzeyde farklılaştığını söyleyemeyiz. Neyse ki deneylerimizdeki sistematik değişkenliğin kaynaklarını daha spesifik bir biçimde belirleyebileceğimiz analiz teknikleri vardır. Önemle tavsiye edilen bir yaklaşım, güven aralıklarının kullanımıdır (bk. Bölüm 11). Güven aralıkları, örneklemlerimizin tahmin ettiği evren ortalamalarına ilişkin örüntüler hakkında kanıt sunabilmektedir (özellikle Kutu 11.5'e bakınız). Bir başka teknik, iki ortalamanın karşılaştırılmasıdır. Öncelikle bağımsız gruplar için ANOVA'ya dair etki büyüklüğü ölçüsünü ve bu desen için güç analizini tartışacağız; sonrasında ise dikkatimizi ortalamaların ikili olarak karşılaştırılmasına yönelteceğiz.

Üç veya Daha Fazla Bağımsız Grup İçeren Desenler İçin Etki Büyüklüğünün Hesaplanması

Psikoloji literatüründe belirli bir araştırma desenine, test istatistiğine ve araştırmanın daha başka özelliklerine bağlı olarak çok farklı etki büyüklüğü ölçülerinin olduğunu belirtmiştik (örneğin; Cohen, 1992; Kirk 1996; Rosenthal ve Rosnow, 1991). Elimizde etki büyüklüğüne ilişkin bir ölçüm olduğu zaman, genellikle bu ölçümü başka bir karşılaştırılabilir etki büyüklüğü ölçümüne pek bir zorluk çekmeden çevirebiliriz. İkiden fazla grubun yer aldığı deneylerde kullanılan önemli bir etki büyüklüğü ölçüsü sınıfı, "ilişkinin gücüne" ilişkin ölçümlere dayanır (Kirk, 1996). Bu ölçümlerin genel bir özelliği, toplam değişkenliğin ne kadarının bağımsız değişkenin bağımlı değişken üzerindeki etkisi tarafından açıklanabildiğini gösteren oranın tahmin edilebilmesini sağlamalarıdır. Popüler bir ilişki gücü ölçüsü **eta kare**, yani η^2'dir. Eta kare, ANOVA Özet Tablosundan (Tablo 12.3) yararlanılarak, genel *F*-testi için kolayca hesaplanabilir (ama birçok bilgisayar programı eta kare değerini etki büyüklüğünün bir ölçüsü olarak otomatik biçimde vermektedir). Eta kare şu formülle belirlenir:

Temel Kavram

$$\frac{\text{Gruplar arası kareler toplamı}}{\text{Toplam kareler toplam}}$$

Örneğimizde ise (bk. Tablo 12.3),

$$\text{eta kare } (\eta^2) = \frac{54.55}{[(54.55) + (37.20)]} = .59$$

Eğer ANOVA tablosu mevcut değilse eta kare, gruplar arası etkiye ait *F*-oranından yararlanılarak da doğrudan hesaplanabilir (bk. Rosenthal ve Rosnow, 1991, s. 441):

$$\text{eta kare } (\eta^2) = \frac{(F)(sd\ etki)}{[(F)(sd\ etki)] + (sd\ hata)}$$

Örneğimizde ise,

$$\text{eta kare } (\eta^2) = \frac{(7.80)(3)}{[(7.80)(3)] + 16} = .59$$

Temel Kavram

Bir başka ölçüm, J. Cohen'in, üç ve daha fazla bağımsız grubu içeren desenler için geliştirdiği *f* dir (bk. Cohen, 1988). Bu değer, etki büyüklüğünün standardize edilmiş bir ölçüsüdür ve *d* ile benzerdir. Fakat iki ortalama arasındaki farka ilişkin etkiyi belirleyen *d*'den farklı olarak, **Cohen** *f* değeri, grup ortalamaları arasındaki yayılışa dayanarak etkiyi belirlemektedir. Hem *d* hem de *f* etkiyi, evren içi standart sapmaya göre (yani ona göre "standardize ederek") ifade etmektedir. Cohen, *f*'in yorumlanması ile ilgili olarak bir yönerge hazırlamıştır. Buna göre, küçük, orta ve büyük etki büyüklükleri, sırasıyla, .10, .25, ve .40 düzeyindeki *f* değerlerine karşılık gelmektedir. *f*'nin hesaplanması, ANOVA Özet Tablosu (Tablo 12.3) kullanılarak kolayca yapılamaz fakat eta kare değeri biliniyorsa, fazla zorluk çekmeden elde edilebilmektedir (bk. Cohen, 1988); şöyle ki:

$$f = \sqrt{\frac{\eta^2}{1 - \eta^2}}$$

Örneğimizde ise,

$$f = \sqrt{\frac{.59}{1 - .59}} = 1.20$$

Böylece, şu sonuca varabiliriz: Bellek eğitimi bağımlı değişkendeki toplam değişkenliğin .59'unu açıkladı ve 1.20 değerinde standart bir *f* etki büyüklüğü ortaya çıkardı. Cohen'in *f* değerini yorumlaman yönergesine göre (.10, .25, .40), bellek eğitimi açık şekilde, anımsama puanları üzerinde büyük bir etki yapmıştır.

Bağımsız Gruplar Deseni İçin Gücün Değerlendirilmesi

Etki büyüklüğü belirlendikten sonra, belirli bir örneklem büyüklüğü için güç tahmini ve F-oranının payıyla (gruplar arası etki) bağlantılı serbestlik derecesi elde edebiliriz Örneğimizde, alfa düzeyini .05 olarak belirledik; deneyde gruplar arası etki için $n = 5$ ve $sd = 3$ (grup sayısı – 1) idi. Veri kümemizin etki büyüklüğü, yani *f* çok yüksektir (1.20); dolayısıyla istatistiksel olarak anlamlı olan bu yüksek düzeydeki etki ile ilgili güç analizi yapmak için geçerli bir neden yoktur.

Bununla birlikte, örneğimizdeki ANOVA'nın anlamlı olmayan bir *F* değeri ile, *f* = .40 gibi bir etki büyüklüğüyle sonuçlandığını varsayın (Cohen'in yönergesine göre bu yine de büyük bir etkidir). Burada yanıtlanması gereken önemli bir soru, "Deneyimizin gücü neydi?" sorusudur. Bu büyüklükteki bir etkiyi, .05 düzeyindeki bir alfa, $n = 5$ büyüklüğündeki bir örneklem ve $sd = 3$ için görme olasılığımız neydi? Güç analizi, bu koşullar altında, gücün .26 olduğu sonucunu verecektir. Başka bir deyişle, bu durumda istatistiksel olarak anlamlı bir sonuca ulaşma olasılığı sadece .26'dır. Bu koşullar altında, yaklaşık olarak denemelerin sadece dörtte birinde istatistiksel olarak anlamlı bir sonuç elde edebiliriz. Bu durumda, deneyin yeterli güce sahip olmadığı sonucu çıkacaktır ve SHAT'nin anlamlı bir sonuç vermemesi üzerinde çok fazla durmak mantıklı değildir. Eğer bunun üzerinde çok durulursa, bağımsız değişkenimizin etkisinin aslında yüksek olduğuna dair önemli gerçek göz ardı edilmiş olur.

Sonuçlar elde edildikten sonra güce ilişkin bilgi edinmek, özellikle SHAT kullanarak anlamlı olmayan bir sonuç elde ettiğimizde, önemli olsa da, ideal olarak güç analizi, istatistiksel olarak anlamlı bir sonuç elde etmenin *apriori*

(başlangıçta) olasılığını saptamak için, deneyden önce yapılmalıdır. Gücün sadece .26 olduğunu bilerek bir deneye başlayan deneyci, anlamlı bir etki elde *etmeme* olasılığı .74 olduğu için, kaynakları ve zamanı boşa harcıyor gibi görünmektedir. Deneyin henüz yapılmadığını ve araştırmacının bellek eğitimi ile ilgili literatürü tarayarak, önceki araştırmacıların bu konuda sıklıkla yüksek bir etki elde ettikleri sonucuna vardığını varsayalım. Ayrıca araştırmacının, deneyin gücünün .80 olmasını istediğini de varsayalım. Güç, örneklem büyüklüğünün artmasıyla artacağı için araştırmacı, .80 düzeyindeki güçle birlikte, yüksek bir etki elde edebilmek için örneklem büyüklüğünün ne kadar olması gerektiğini bulmak isteyecektir. Güç analizi bu işi yapabilir ve araştırmacı deneye başlamadan önce bu bilgiyi dikkate almalıdır.

Çoklu Grup Deneylerinde Ortalamaların Karşılaştırılması

Daha önce belirttiğimiz gibi, tek faktörlü ve çok gruplu bir deneyde "bir şeylerin gerçekleştiğini" biliyor olmak, genellikle çok ilgi çekici değildir. Çoğunlukla, zihnimizdeki "bu değişkenin bir etkisi vardır"dan daha spesifik bir hipotezle hareket ederek araştırma yaparız ya da en azından yapmalıyız. Ne genel *F* sonuçları, ne de genel etki büyüklüğü ölçüsü hangi ortalamaların hangi ortalamalardan farklılaştığını bize söyleyemez. Örneğin, bellek deneyimizdeki dört ortalamaya bakıp "imgeleme" ortalamasının, "hikâye" ortalamasından anlamlı derecede farklılaştığına karar veremeyiz. Buradaki genel *F* sonuçları bize tüm gruplar arasında, şans eseri beklenebilecek değişkenlikten daha yüksek düzeyde bir değişkenliğin var olduğunu söylemektedir.

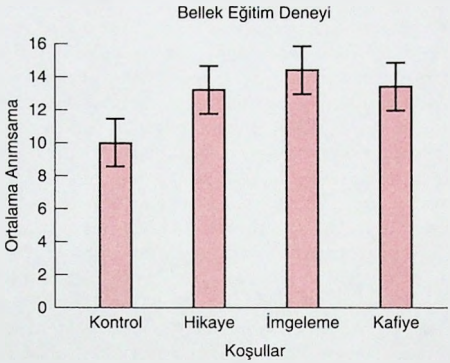
Tek faktörlü ve çok gruplu bir deneyde neyin gerçekleşmiş olduğunu öğrenebilmek için birbirini tamamlar nitelikte iki yol önerebiliriz. Bir yaklaşım, deneyimizdeki ortalama tahminlerine ait %95'lik güven aralıklarını hesaplayarak, evren ortalamalarının olası örüntülerini incelemektir. Bu yaklaşım Bölüm 11'de bir çoklu grup deneyinde ortalamaları karşılaştırmak için güven aralıklarının nasıl kullanılabileceğini ele aldığımızda gösterilmişti. Güven aralıkları, deney gruplarımızın ortalamaları tarafından tahmin edilen evren ortalamaları arasındaki olası farklılıklara ilişkin karara varmak için kullanılabilir. Bu kararlar, güven aralıklarının çakışıp çakışmadığı incelenerek verilmektedir (özellikle bk. Kutu 11.5). Güven aralıklarının genişliğinin, tahminlerimizin kesinliği hakkında bilgi verdiğini hatırlayınız.

Bellek deneyi için güven aralıklarının yapılandırılması, Bölüm 11'de özetlenen işlemlerle yapılır. ANOVA Özet Tablosundaki OK_{hata}'nın karekökü $s_{birleştirilmiş}$'e eşit olduğu için, %95'lik güven aralığını şu şekilde belirleyebiliriz:

$$\%95\ GA = \overline{X} \pm [\sqrt{(OK_{hata}/n)}](t_{kritik}tt)$$

Buradaki t_{kritik}, OK_{hata} ile ilişkilendirilmiş serbestlik derecesi ile birlikte *t* değeridir.

ŞEKİL 12.1 Bellek eğitimi deneyi için ortalamalar ve %95'lik güven aralıkları



Örneğimizde, OK_{hata} için serbestlik derecesi 16'dır (bk. Özet ANOVA tablosu) ve .05 düzeyi için (iki uçlu test) t_{kritik} 2.12'dir. Buradan hareketle,

$$\%95\ GA = \overline{X} \pm \left[\sqrt{(2.33/5)}\right](2.12) = \overline{X} \pm (\sqrt{.466})(2.12) = \overline{X} \pm (.683)(2.12)$$

$$= \overline{X} \pm 1.45$$

Bellek deneyi için, güven aralıklarının yapılandırılmasına dayanan sonuçlar, Şekil 12.1'de gösterilmiştir. Şu anda bu sonuçları yorumlayabiliyor olmalısınız fakat ilgili bilgileri hatırlamak isterseniz Kutu 11.5'e bakın.

Temel Kavram

İkinci bir yaklaşım SHAT'den yararlanmakta ve bağımsız değişkenimizin genel etkisinin kaynağını belirlemek için ikili grup karşılaştırmalarına odaklanmaktadır. **İki ortalamanın karşılaştırılması** araştırmacıya, ilgilenilen belirli bir farka odaklanma olanağı sağlar. İki ortalamanın karşılaştırılması epey karmaşık olabilir; örneğin, bir deneyde iki veya daha fazla grubun ortalamasının, bir başka grubun ortalamasıyla ya da başka iki veya daha fazla grubun ortalamasıyla karşılaştırılmasında olduğu gibi. Bununla birlikte, çoğunlukla sadece ayrı gruplara ait iki ortalamanın karşılaştırılmasıyla ilgileniriz. Bu türden ikili ortalama karşılaştırmaları genellikle genel *F*-testinin istatistiksel olarak anlamlı olduğunu belirlediğimizde yapılır.

Ortalamaların ikili karşılaştırmalarını yapmanın bir yolu *t*-testi kullanmaktır fakat çoklu grup deneylerinde ortalamalar karşılaştırılırken *t*'nin hesaplanmasında küçük bir değişiklik yapılır. Buna göre, genel *F*-testimizdeki grup içi değişkenlik tahminini temel alarak (OK_{hata}), birleştirilmiş bir değişkenlik tahmini kullanmak isteriz. Başka bir deyişle, değişkenlik tahminimiz, sadece ilgilenilen iki gruba ilişkin bilgiyi değil, deneydeki tüm gruplardan elde edilen bilgiyi kullanır. Bu nedenle bu tür bir *t*-testinin formülü şu şekilde olacaktır: tahminimiz,

$$t = \frac{\overline{X}_1 - \overline{X}_2}{\sqrt{[OK_{hata}]\left[\frac{1}{n_1} + \frac{1}{n_2}\right]}}$$

sadece ilgilenilen iki gruba ilişkin bilgiyi değil, deneydeki tüm gruplardan elde edilen bilgiyi kullanmaktadır. OK_{hata}'nın değeri, genel F-testimize ait özet ANOVA tablosundan elde edilmiştir ve karşılaştırma için kullandığımız t-testleri için serbestlik derecesi, OK_{hata}'ya ilişkin olan serbestlik derecesi ile aynıdır [yani $k(n - 1)$, burada k = grup sayısı]. Örneğin, Tablo 12.3'te rapor edilen analiz için denek içi (hata) OK, 16 serbestlik derecesi ile birlikte, 2.33'tür [$4(5 - 1) = 16$].

Bellek deneyi için yapabileceğimiz ikili ortalama karşılaştırmalarından biri, bellek eğitimi gruplarının (birleştirilmiş) ortalama performansını, kontrol grubu ile karşılaştırmak olabilir. Üç bellek eğitimi grubu için ortalama bellekte tutma miktarı 13.67 (n = 15) ve kontrol grubu için ortalama 10.00'dır (n = 5). Böyle bir karşılaştırmayla "Eğitim türünden (öykü, imgeleme, kafiye) bağımsız olarak, bellek eğitimi, eğitimin olmadığı duruma göre (kontrol), bellekte daha etkin bir bellekte tutma sürecine neden olabilir mi?" sorusunu sorabiliriz. Sıfır hipotezi evren ortalamalarının farklılaşmadığı yönündedir (örneklem ortalamaları ise sadece tesadüf eseri farklılaşmaktadır). Gerekli olan değerler, yukarıda t için verilen formülde yerine konduğunda, istatistiksel olarak anlamlı bir etki gözleriz, $t(16) = 4.66$, $p = .0003$. Bu nedenle bu deneyde, eğitim türünden bağımsız olarak, bellek eğitimi, eğitimin yapılmadığı duruma kıyasla, sözcüklerin bellekte daha iyi tutulmasını sağlamıştır. Açıkça görebilirsiniz ki bu ifade, sadece deneyin dört koşulu arasındaki değişkenliğin, tesadüf eseri beklenecek değişkenlikten daha büyük olduğunu söyleyen genel F-testinin ifadesinden daha spesifiktir.

Ortalamaların ikili olarak karşılaştırılmasına ilişkin Cohen d, t-testinin sonuçları kullanılarak hesaplanabilir. Buradaki duruma ait Cohen d formülü şöyledir:

$$d = \frac{2(t)}{\sqrt{sd_{hata}}}$$

Üç bellek eğitimi grubunun ve kontrol grubunun karşılaştırılması amacıyla, 4.66 değeri, 16 sd_{hata} için formülde yerine konduğunda, etki büyüklüğü olan d, 2.33 olarak hesaplanmaktadır. Cohen'nin etki büyüklüklerine ait kriterine göre bu değer, bellek yönergesi için, yönergenin olmadığı duruma göre yüksek bir etki olarak yorumlanabilir.

t-testini kullanırken, sıfır hipotezini belirli bir olasılık dâhilinde (örneğin, p = .05) reddetmek veya reddetmemek üzere bir karar vermeye çalışırız. Önceden de belirtildiği gibi, SHAT'nin sonucu ile ilgili tam olasılık, sonuçları yorumlarken önemli olabilir (örneğin, Posavac, 2002). Tam olasılık düştükçe, çalışmanın tam bir tekrarının, sıfır hipotezini $p < .05$ düzeyinde reddetmeye izin verme olasılığı artar (bk. Zechmeister ve Posavaz, 2003). Asgari düzeyde, istatistiksel anlamlılık için bilgimiz olan en düşük olasılığı rapor etmek isteriz. (Bilgisayarlar otomatik olarak test sonucumuzun tam olasılığını verir.)

t karşılaştırmasının sonuçları ayrıca, sonuçları önceki çalışmalarla iki şekilde karşılaştırmamızı sağlar. Öncelikle, deneyimizin istatistiksel anlamlılık sonuçlarının önceki deneylerde gözlenenlerle aynı olup olmadığını ele alabiliriz. Başka bir deyişle, istatistiksel olarak anlamlı bir sonucu *tekrarladık* mı? İkinci olarak, bu ikili ortalama karşılaştırması için, önceki deneylerden elde edilen etkilerle karşılaştırılabilecek olan (muhtemelen bir meta-analizin parçası

olarak) etki büyüklüğünü hesaplayabiliriz (örneğin, Cohen *d*). Bu iki karşılaştırmanın hiçbirini güven aralıklarını kullanarak yapmak kolay değildir. Şöyle ki, SHAT'den farklı olarak güven aralıkları, deneyimizde görülen farka ilişkin tam olasılığı vermez ve *t*-testinin ardından, etki büyüklüklerinin hesaplanması, daha doğrudan bir şekilde yapılmaktadır (bk. Bölüm 11).

Özetle, verinizi ve ortalamalar arası farkları birden fazla istatistiksel tekniği kullanarak incelemenizi, "neyin gerçekleştiği" ile ilgili kanıtı, veri analizine dair farklı yaklaşımlardan hareketle aramanızı öneriyoruz.

TEKRARLI ÖLÇÜMLER VARYANS ANALİZİ

- Tekrarlı ölçümler varyans analizini kullanarak sıfır hipotezinin test edilmesine ait genel işlemler ve bu testin altında yatan mantık, bağımsız gruplar varyans analizi için kullanılanlar ile aynıdır.
- Tamamlanmış bir tekrarlı ölçümler deseni için varyans analizine başlamadan önce her katılımcı için ve her koşula ait bir özet puan (örneğin; ortalama, ortanca) hesaplanmalıdır.
- Tüm katılımcılar arasında, bağımsız değişkenin her koşuluna ait performansı özetlemek amacıyla betimsel veriler hesaplanır.
- Tekrarlı ölçümler deseninde varyans analizinin başlıca farkı, hata değişkenliğinin, yani artık değişkenliğin nasıl tahmin edildiği ile ilgilidir; artık değişkenlik, bağımsız değişkenden ve deneklerden kaynaklanan sistematik değişkenlik, toplam değişkenlikten çıkarıldığında kalan değişkenliktir.

Tekrarlı ölçümler deseninde deneylerin analiz işlemleri, bağımsız gruplar deseni kullanan deneylerdekiyle aynıdır. Deneyde elde edilen farkların, tek başına hata değişkenliğinden beklenen farklardan daha büyük olup olmadığını belirlemek için SHAT'nin ilkeleri uygulanır. Analiz, bağımsız değişkenin, bağımsız değişken düzeyleri arasında sistematik bir değişkenlik üretip üretemediğini belirlemek amacıyla, genel varyans analizi ile başlar. Bu genel analizin istatistiksel olarak anlamlı sonuç vermesi şunu kanıtlar: Güven aralıkları ve ortalamaların ikili karşılaştırılması, sistematik değişkenliğin özgül kaynağının bulunması amacıyla kullanılabilir –başka bir deyişle, hangi düzeylerin kendi aralarında farklılaştığını saptamak için kullanılabilir. Bağımsız gruplar deseni kullanan deneyler için bu genel analiz planının altında yatan mantığı ve bu plana ilişkin genel işlemleri daha önce açıklamıştık. Bu bölümde, tekrarlı ölçümler desenlerine özgü analiz özelliklerine odaklanacağız ve bir ANOVA Özet Tablosu örneğini açıklayacağız. Bu analizin bir gösterimini yapmak için kullanılan veri, Bölüm 7'de açıklanan zaman algısı deneyine dayanmaktadır ve devam etmeden önce oradaki tartışmayı gözden geçirmek isteyebilirsiniz.

Verileri Özetleme Bir tekrarlı ölçümler deseninde, her katılımcının, deneyin her koşulunda yer aldığını hatırlayın. Tamamlanmış bir desende, her katılımcı, her koşulu birden fazla sayıda alır; tamamlanmamış bir desende ise her katılımcı, her koşulu tam olarak bir kez alır. Bölüm 7'de, tamamlanmış bir tekrarlı

TABLO 12.4 BİR TEKRARLI ÖLÇÜMLER DESENİ DENEYİ İÇİN VERİ MATRİSİ VE VARYANS ANALİZİ ÖZET TABLOSU

Veri Matrisi

	Aralık Uzunluğu			
Katılımcı	12	24	36	48
1	13	21	30	38
2	10	15	38	35
3	12	23	31	32
4	12	15	22	32
5	16	36	69	60
Ortalama (S)	12.6 (2.0)	22.0 (7.7)	38.0 (16.3)	39.4 (10.5)

Not: Tablodaki her değer, katılımcıların, aralık uzunluğu değişkeninin her düzeyindeki ortancalarını göstermektedir.

Değişkenliğin Kaynağı	*sd*	*KT*	*OK*	*F*	*p*
Denekler	4	1553.5	—	—	—
Aralık uzunluğu	3	2515.6	838.5	15.6	.000
Artık (hata değişkenliği)	12	646.9	53.9		
Toplam	19	4716.0			

ölçümler deseninde, katılımcıların dört zaman aralığı (12, 24, 36 ve 48 saniye) için tahmin yaptıkları bir deney açıklamıştık. Örneğin, tek bir denemede katılımcılara, seçkisiz olarak belirlenmiş bir zaman aralığı (örneğin, 36 saniye) verildi ve bu aralığın süresini tahmin etmeleri istendi.

İlk adım, her bir bireyin her koşuldaki performansını özetlemek amacıyla bir puanın hesaplanmasıdır. Zaman algısı deneyinde katılımcılara her koşul altı kez verildi; bu nedenle, deneydeki dört koşul için her katılımcı 24 tahmin yaptı. Her bir katılımcının dört koşulun her birindeki performansını özetlemek için ortanca kullanıldı. Verilerin özetlenmesinde bir sonraki adım, tüm katılımcılar arasında, her bir koşula ait betimsel istatistiklerin hesaplanmasıdır. Her koşula ait ortalamalar ve standart sapmalar (parantez içinde) Tablo 12.4'tedir (Ayrıca Tablo 7.3'e bakınız).

Buradaki analizin odak noktası, katılımcıların farklı uzunluktaki zaman aralıklarını ayırt edip edemeyeceğidir. Tablo 12.4'te yer alan verilerin genel varyans analizine ait sıfır hipotezi, her aralık için tahmin edilen evren ortalamalarının aynı olduğu şeklindedir. Bu sıfır hipotezine ait *F*-testini yapabilmek için hata değişkenliğinin ve sistematik değişkenliğin toplamına (*F*-testinin payı) ait bir tahmine ihtiyacımız vardır. Dört farklı aralık için katılımcılar arasındaki ortalama tahminleri, pay için ihtiyacımız olan bilgiyi vermektedir. Eğer farklı aralık uzunlukları katılımcıların yargılarını sistematik şekilde etkilediyse, aralıklar için ortalamatahminlerinin bu sistematik değişkenliği yansıtacağını biliyoruz. *F*-testini tamamlamak için, sadece hata değişkenliğinin de (*F*-testinin paydası)

KUTU 12.4

TEKRARLI ÖLÇÜMLER DESENİNDE HATANIN VE DUYARLILIĞIN TAHMİN EDİLMESİ

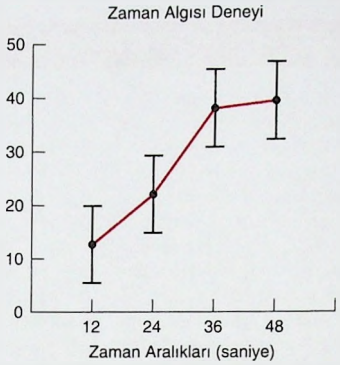
Tekrarlı ölçümler desenlerinin ayırıcı bir özelliği, hata değişkenliğinin tahmin edilme şekliyle ilgilidir. Seçkisiz gruplar deseni için, farklı gruplarda dengelenen katılımcılar arasındaki bireysel farklılıkların, *F*-testinin paydası olan hata değişkenliğinin bir tahminini verdiğini daha önce açıklamıştık. Bu desenlerde bireyler sadece bir koşula katıldıkları için, katılımcılar arasındaki farklar ortadan kaldırılamaz – sadece dengelenebilir. Öte yandan, tekrarlı ölçümler desenlerinde, katılımcılar arasında sistematik bir değişkenlik mevcuttur. Bazı katılımcılar koşullar arasında tutarlı şekilde diğerlerine göre daha iyi bir performans gösterir, bazıları ise tutarlı şekilde daha kötü performans sergiler. Bununla birlikte, her bir birey tekrarlı ölçümler deseninin her koşuluna katıldığı için, katılımcılar arasındaki farklar, her koşuldaki ortalama performansa eşit miktarda katkı yapar. Bu nedenle, tekrarlı ölçümler desenlerinde her bir koşul arasındaki herhangi bir fark, katılımcılar arasındaki sistematik farkların bir sonucu olamaz. Bununla birlikte tekrarlı ölçümler desenlerinde katılımcılar arasındaki farklar sadece dengelenmekle kalmaz, analize olan etkisi gerçekten ortadan kalkar. *Tekrarlı ölçümler desenlerinin katılımcılardan kaynaklanan sistematik değişkenliği ortadan kaldırabilme özelliği, bu desenleri genel olarak seçkisiz gruplar desenlerinden daha duyarlı kılmaktadır.*

bir tahminine ihtiyacımız vardır. Tekrarlı desendeki değişkenliğin kaynağı, koşulların farklı katılımcıları etkileme biçimlerindeki farklılıklarda yer almaktadır. Bu değişkenlik tahminine, *artık değişkenlik* (residual variation) denmektedir. Kutu 12.4'e bakınız.

ANOVA Özet Tablosunun Yorumlanması Bu analiz için varyans analizi özet tablosu Tablo 12.4'ün alt kısmında sunulmuştur. Tekrarlı ölçümler varyans analizinin hesaplamaları aslında bilgisayarda istatistiksel paket program kullanılarak yapılmaktadır. Bu nedenle şimdiki odak noktamız, özet tablosundaki değerlerin nasıl hesaplandığı değil, nasıl yorumlandığıdır. Tablo 12.4, değişimlenen bir bağımsız değişkenin yer aldığı tekrarlı ölçümler deseni analizindeki değişkenliğin dört kaynağını veriyor. Tablonun altından üstüne doğru bu kaynaklar; (1) toplam değişkenlik, (2) artık değişkenlik, (3) aralık uzunluğundan kaynaklanan değişkenlik (bağımsız değişken) ve (4) deneklerden kaynaklanan değişkenliktir.

Herhangi bir özet tablosunda olduğu gibi en kritik bilgiler, ilgilenilen bağımsız değişkenin etkisine ait *F*-testi ve sıfır hipotezinin doğru olduğu varsayımı altında bu *F*-testi ile ilişkilendirilen olasılık değeridir. Tablo 12.4'teki önemli *F*-testi, aralık uzunluğuna ait olan *F*-testidir. Bu *F*-testi için pay, aralık uzunluğuna ait ortalama karedir (*OK*); payda ise artık *OK*'dir. Dört aralık uzunluğu, dolayısıyla pay için 3 serbestlik derecesi (*sd*) vardır. Artık değişkenlik için 12 *sd* söz konusudur. Artık değişkenliğine ait *sd*'yi, deneklere ait ve aralık uzunluğuna ait *sd*'leri toplam *sd*'den çıkararak elde edebiliriz (19 – 4 – 3 = 12). Elde edilen 15.6 değerindeki *F*, sıfır hipotezi altında, .0004 düzeyinde bir olasılığa sahiptir ve bu olasılık, istatistiksel anlamlılık kriterimiz için seçtiğimiz .05 anlamlık düzeyinden küçüktür. Buradan hareketle sıfır hipotezini reddederiz ve

ŞEKİL 12.2 Zaman algısı deneyine ait ortalamalar ve %95'lik güven aralıkları



aralık uzunluğunun, sistematik değişkenliğin kaynağı olduğu sonucuna varırız. Bu, katılımcıların tahminlerinin, aralık uzunluğunun bir fonksiyonu olarak, sistematik şekilde değiştiği anlamına gelir.

Şekil 12.2, zaman algısı deneyinde yer alan ortalamalar etrafındaki %95'lik güven aralıklarını göstermektedir. Bu aralıkların yapılandırılmasına ait işlem, bağımsız gruplar deneyindeki işlemle aynıdır. Aralıklar, genel ANOVA'daki OK_{hata}'nın (artık) kullanılmasıyla yapılandırılmıştır [Loftus ve Masson'un (1994) önerdiği gibi]. Şöyle ki,

$$\%95\ GA = \overline{X} \pm \left[\sqrt{(OK_{hata}/n)}\right](t_{kritik})$$

buradaki t_{kritik}, OK_{hata} (artık) ile ilişkilendirilen serbestlik derecesine ait t değeridir. Tekrarlı ölçümler desenlerinde güven aralıklarının yorumlanması, bağımsız gruplar desenine ait olanla aynıdır (bk. Bölüm 11).

Etki Büyüklüğü Ölçüleri Önceden değindiğimiz gibi, analizinize etki büyüklüklerini dâhil etmek iyi bir fikirdir. Tekrarlı ölçümler deseni için etki büyüklüğünün tipik bir ölçüsü, eta kare (η^2) adı verilen ilişkinin gücüdür. Denek içi etkinin kareler toplamı, denek içi etki ile artık, yani hatanın birleştirilmiş kareler toplamına bölünmesiyle elde edilebilir. Örneğimiz için,

$$\text{eta kare } (\eta^2) = \frac{KT_{hata}}{KT_{hata} + KT_{hata}} = \frac{2515.6}{2515.6 + 646.9} = .795$$

Bu, bağımsız değişkenin sorumlu olduğu değişkenliğin oranını gösterir. Bazı durumlarda genel varyans analizinden sonra, bağımsız gruplar deseninde gördüğümüz gibi, ikili ortalama karşılaştırmaları gelir.

BAĞIMSIZ GRUPLAR DESENLERİ İÇİN İKİ FAKTÖRLÜ VARYANS ANALİZİ

Bağımsız gruplar desenleri için iki faktörlü varyans analizi, iki bağımsız değişkenin her birinin, iki veya daha fazla düzeyde değişimlendiği deneylerin analizinde kullanılır. İki bağımsız değişkenli karmaşık desenlerin altında yatan mantık ve bu deneylerin analizine ait kavramsal temel, Bölüm 8'de açıklanmıştır. Ayrıca, temel etkileri ve etkileşimleri betimlemeyi de öğrendiniz. Bu bölümde faktöryel bir desenin, A'nın temel etkisi, B'nin temel etkisi ve A × B etkileşimi için *F*-testlerini içeren bilgisayar destekli analizine odaklanacağız. Bağımsız gruplar için iki faktörlü analiz, iki bağımsız değişkenin de seçkisiz gruplar deseni kullanılarak değişimlendiği, iki bağımsız değişkenin de doğal gruplar desenini temsil ettiği ya da bir bağımsız değişkenin doğal gruplar desenini, diğerinin ise seçkisiz gruplar desenini temsil ettiği deneylere uygulanabilmektedir. Bölüm 8'de değindiğimiz gibi, karmaşık bir desenin analizi, genel *F*-testinin, bir etkileşimin varlığını gösterdiği ve göstermediği durumlarda biraz farklı ilerlemektedir. Önce, bir etkileşimin saptandığı durumdaki analiz planını ele alacağız.

Etkileşimin Olduğu Karmaşık Desen Analizi

- Eğer genel varyans analizi istatistiksel olarak anlamlı bir etkileşimin varlığını göstermişse, etkileşimin kaynağı, basit temel etki analizleri ve ikili ortalama karşılaştırmaları kullanılarak belirlenir.
- Basit temel etki, bir bağımsız değişkenin, ikinci bağımsız değişkenin bir düzeyindeki etkisidir.
- Eğer bir bağımsız değişkenin üç veya daha fazla düzeyi varsa, ortalamaları ikişerli karşılaştırarak, basit temel etkinin kaynağını araştırmak için ikili ortalama karşılaştırmaları kullanılabilir.
- Evren ortalamalarının tahminine ait kesinlik hakkında bilgi vermek amacıyla, grup ortalamaları etrafında güven aralıkları çizilebilir.

İki bağımsız değişken (A × B) içeren, her bir bağımsız değişkenin bağımsız gruplarının olduğu (seçkisiz gruplar veya doğal gruplar) hipotetik bir karmaşık desen düşünün. Değişken A'nın iki düzeyi ve Değişken B'nin üç düzeyi vardır. Dolayısıyla desen bir 2 × 3 bağımsız gruplar desenidir. Deneyin ayrıntıları bizi ilgilendirmemekle birlikte, her grupta beş katılımcının yer aldığını varsayalım (n = 5; N = 30). Tablo 12.5, bu örnekteki altı grup için ortalama performansı gösteriyor. Şu anda, veride ne gibi eğilimlerin bulunduğunu görmek için öncelikle özet istatistiklerin incelendiğini biliyor olmanıza güveniyoruz. Tablo 12.5'te görülen sonuçları nasıl betimlersiniz? Sonuçları betimlemenin bir yolu, A'nın ilk düzeyi için, yani a_1 için, B'nin üç düzeyi arasında çok küçük bir farklılığın bulunduğunu ifade etmektir. Öte yandan, a_2 düzeyi için, B'nin aynı üç düzeyi arasında ortalamalar oldukça değişmiştir. Aynı zamanda bu sonuçları betimlemenin başka bir yolu, Bölüm 8'de, karmaşık desenlere ilk olarak değinildiğinde tartıştığımız çıkarma yöntemini kullanmaktır. Bu yöntemi kullanmak, bir etkileşimin olup

TABLO 12.5 HİPOTETİK BİR 2 × 3 DESENDE GRUPLARIN ORTALAMA PERFORMANSI

		Değişken B		
		b_1	b_2	b_3
Değişken A	a_1	19.0	19.0	20.0
	a_2	10.6	15.8	18.2

olmadığını belirlemede yardımcı olacaktır. B'nin her düzeyinde (b_1, b_2, b_3), a_1 ve a_2 için, iki ortalama arasındaki farkı incelemek, üç farkın (8.4, 3.2, 1.8), birbirinden farklı olduğunu gösterecektir. Bu durum, bir etkileşimin varlığını işaret etmektedir. Bölüm 8'de öğrendiğiniz gibi, ortalamaların bir grafiğini çıkarmak da bu etkileşimin niteliğini görmemize yardımcı olmaktadır. Genel *F*-testinin, etkileşimin istatistiksel olarak anlamlı olduğunu ($p < .05$) onayladığını varsayalım.

İki bağımsız değişkenin bir etkileşimi olduğunu onayladığımızda, söz konusu etkileşimin kaynağını daha hassas bir şekilde belirlemeliyiz. Özellikle, anlamlı bir etkileşimin kaynağının izini sürmek için tasarlanmış olan istatistiksel testler mevcuttur. Bu testler basit temel etkiler ve ortalamaların ikili karşılaştırmaları adını alır (bk. Keppel, 1991) ve Bölüm 8'de kısaca tartışılmıştır. İki ortalama arasındaki karşılaştırmalara bu bölümde de daha önce değinildi.

Basit temel etkinin, bir bağımsız değişkenin, ikinci bağımsız değişkenin bir düzeyindeki etkisi olduğunu hatırlayınız. Aslına bakılırsa etkileşimin bir tanımı, basit temel etkilerin düzeyler arasında farklılaşmasıdır. 2 x 3'lük bir desende aslında beş basit temel etki vardır. Basit temel etkilerin üçünü, Değişken A'nın, Değişken B'nin her bir düzeyindeki etkisi temsil etmektedir. Diğer iki basit temel etkiyi, Değişken B'nin, Değişken A'nın her bir düzeyindeki etkisi temsil etmektedir. Analiz için hangi basit temel etki kümesinin seçileceği, deneyin altında yatan gerekçeye bağlıdır. Başka bir deyişle, sonuçların yorumlanması için belirli bir basit temel etki grubunun vurgulanması, bir diğerinin vurgulanmasından daha önemli olabilmektedir. Tabi her iki değişkenin düzeyleri için basit temel etkilerin farklı olduğunu bulmak, bir etkileşimin varlığını gösterir.

Bir basit temel etkiyi nasıl hesaplarız? İstatistik paket programları hesaplanacak olan basit temel etki analizlerine her zaman izin vermez ve izin verdiğinde de izlenen belirli sayısal işlemler farklılaşabilir. Bu analizleri hesap makinasıyla yapmak için göreli olarak basit yollar vardır (Zechmeister ve Posavac, 2003). Ama biz, ANOVA paket programı kullanarak kolayca yapılabilen aşağıdaki işlemi önerebiliriz.

Yukarıdaki örneğimizi ele alın. Değişken A'nın birinci düzeyi olan a_1 için basit temel etkiyi analiz etmek istediğimizi varsayın. Bu analizde üç "grup" yer almaktadır (a_1b_1, a_1b_2, a_1b_3). Bir yaklaşım, bu veride bağımsız gruplar için basit (tek yönlü) ANOVA uygulamaktır. Başka bir deyişle, bir bağımsız değişkenin üç düzeyine atanan ve üç seçkisiz gruptan oluşan katılımcılar söz konusudur. Bu analizi yapın ve ANOVA Özet Tablosu'ndaki gruplar arası ortalama kareyi (*OK*) belirleyin (yani, değişkeninizin etkisi için olan *OK*). Bu, gruplar arası kareler toplamının, buna ait *sd*'ye, yani grup sayısı eksi 1'e bölünmüş hâlidir ve burada 3 – 1, *sd* = 2'dir. *F*-oranını elde edebilmek için bu analizdeki

gruplar arası *OK*'yi, 2 x 3'lük karmaşık desende, etkileri incelemek için yaptığınız genel *F*-testine ait OK_{hata}'ya (denek içi) bölmeniz gerekecektir. Örneğimizdeki 2 × 3'lük desende, 30 katılımcı için OK_{hata}'ya ait *sd*, 24'tür, dolayısıyla kritik *F* değeri 2 ve 24 serbestlik dereceleri ile ilişkilendirilecektir.

Hipotetik deneyimizdeki basit temel etkilerin ikisi, üç ortalamayı içeriyor (yani, B'nin üç düzeyindeki, a_1 ve a_2 düzeylerini). Eğer istatistiksel analiz bu düzeylerin birinde anlamlı bir basit temel etki ortaya çıkarırsa, ortalamalar arasında fark olduğu sonucu çıkarılabilir (yani Değişken A'nın belirli bir düzeyindeki üç ortalama arasında). Bu durumda bir sonraki adım, basit temel etkinin tümüyle analiz edilebilmesi için ikili ortalamaların karşılaştırılmasıdır. İkili ortalamaların karşılaştırılması, düzeyler arasındaki farklılıkların niteliğini belirlemeye yardımcı olacaktır. İki ortalama arasındaki karşılaştırmanın analizinde, bu bölümde daha önce değindiğimiz *t*-testi kullanılır. .05 düzeyi için kritik *t* değerini bulmak amacıyla, *t* formülünde genel 2 × 3 ANOVA Özet Tablosu'ndaki OK_{hata} ve bu terimle ilişkilendirilen *sd* (örneğimizde 24) kullanılır.

B'nin üç düzeyi için, a_1 ve a_2'deki ortalama performansın karşılaştırılmasında olduğu gibi, bir bağımsız değişkenin sadece iki düzeyi için basit temel etki analizi yapıyorsanız, iki ortalamanın karşılaştırılmasında yaptığınız gibi, *t*-testini kullanabilirsiniz. İki grup için *t*-testindeki örneklem büyüklüklerinin, karşılaştırdığınız iki hücrenin her birindeki katılımcı sayısını temel aldığına dikkat edin. Hipotetik deneyimizde, her grup için $n = 5$'tir. Son olarak, yukarıda tartıştığımız ikili ortalama karşılaştırmalarında yaptığımız gibi, *t*-testinizin hata terimi için 2 × 3 ANOVA'daki OK_{hata}'yı kullanabilirsiniz. Buradaki iki gruplu *t*-testi için olan serbestlik derecesi, genel ANOVA'nızda yer alan OK_{hata} ile ilişkilendirilen serbestlik derecesi olacaktır. İki düzey varsa, basit temel etki iki ortalama arasındaki farkı karşılaştırmaktadır ve ek bir karşılaştırmaya gerek yoktur.

Etkileşim dikkatle analiz edildikten sonra, araştırmacılar ayrıca her bağımsız değişkenin temel etkisini de analiz edebilirler. Bununla birlikte, eğer etkileşim istatistiksel olarak anlamlı ise, temel etkiler genellikle daha az ilgi çekicidir.

Etkileşimin Olmadığı Durumda Analiz

- Genel varyans analizi, bağımsız değişkenler arasındaki etkileşimin istatistiksel olarak anlamlı olmadığını gösterirse, bir sonraki adım, bağımsız değişkenlerin temel etkisinin istatistiksel olarak anlamlı olup olmadığını belirlemektir.
- İstatistiksel olarak anlamlı bir temel etkinin kaynağı, ortalamaların ikişerli karşılaştırmalarını yaparak ve güven aralıklarını yapılandırarak daha hassas şekilde belirlenebilir.

Etkileşim istatistiksel olarak anlamlı değilse bir sonraki adım, her bağımsız değişkenin temel etkisinin incelenmesidir. Bir bağımsız değişkene ait genel temel etki istatistiksel olarak anlamlı değilse, daha fazla yapılacak bir işlem yoktur. Bununla birlikte, bir temel etki istatistiksel olarak anlamlı ise, araştırmacının benimseyebileceği çeşitli yaklaşımlar vardır. Örneğin bir bağımsız değişkenin üç veya daha fazla düzeyi varsa, istatistiksel olarak anlamlı bir temel

etkinin kaynağı, *t*-testlerinin kullanılmasıyla ikili ortalama karşılaştırmaları yapılarak daha hassas şekilde belirlenebilir. Bir diğer yaklaşım, Bölüm 11'de tek faktörlü bağımsız gruplar deseninin analizinde gösterdiğimiz gibi, yine grup ortalamaları etrafında güven aralıklarının yapılandırılmasıdır. Karmaşık desenlerin farkı, bir bağımsız değişkene ait verinin, diğer bağımsız değişkenlerin düzeylerinde birleştirilmesidir.

Bağımsız Grupları İçeren İki Faktörlü Desene Ait Etki Büyüklükleri

ANOVA'nın kullanıldığı bir karmaşık desen için etki büyüklüğünün yaygın bir ölçüsü, daha önce tek faktörlü desenler bağlamında tartışılan eta kare (η^2) veya açıklanan değişkenliğin oranıdır. Eta karenin hesaplanmasında sadece ilgilenilen etkiye odaklanmamız önerilmektedir (bk. Rosenthal ve Rosnow, 1991). Spesifik olarak eta kare aşağıdaki şekilde belirlenebilir:

$$\eta^2 = \frac{KT_{\text{incelenen etki}}}{KT_{\text{incelenen etki}} + KT_{\text{denek içi}}}$$ (bk. Rosenthal ve Rosnow, 1991, s. 352)

Buradan hareketle eta kare, bir A × B desenindeki her üç etki için de hesaplanabilir.

Yukarıda bahsedildiği gibi (ve bakınız, Rosenthal ve Rosnow, 1991), etkilere ait kareler toplamı değerleri elde mevcut değilse eta kare, *F* oranı (ve *sd*) kullanılarak, ilgilenilen her etki için hesaplanabilir.

KARMAŞIK DESENLERİN ANALİZİNDE GÜVEN ARALIKLARININ ROLÜ

Karmaşık bir desenin analizinde, ilgilenilen ortalamalar için güven aralıklarının yapılandırılmasından yararlanılabilir. Örneğin 2 x 3'lük bir desendeki her bir ortalama, Bölüm 11'de ve bu bölümde daha önce özetlenen işlemler sonrasında bir güven aralığıyla sınırlandırılabilir. Formülün aşağıdaki gibi olduğunu hatırlayınız:

%95'lik güven aralığının üst sınırı: $\overline{X} + [t_{.05}][s_{\overline{X}}]$

%95'lik güven aralığının alt sınırı: $\overline{X} - [t_{.05}][s_{\overline{X}}]$

Örneklem büyüklükleri eşit ise, tahminî standart hata şu şekilde belirlenir:

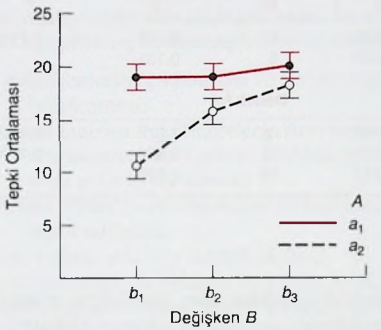
$$s_{\overline{X}} = \frac{s_{\text{birleşik}}}{\sqrt{n}}$$ buradaki n = her bir gruptaki örneklem büyüklüğüdür

ANOVA Özet Tablosu'ndaki OK_{hata}'nın kare kökü $s_{\text{birleşik}}$'e eşit olduğu için, %95'lik güven aralığını şu şekilde belirleyebiliriz:

$$\%95\ GA = \overline{X} \pm (t_{.05})\left[\sqrt{(OK_{\text{hata}}}/\sqrt{n})\right]$$

Buradaki $t_{.05}$, OK_{hata} ile ilişkilendirilen serbestlik derecesi ile belirlenmektedir.

ŞEKİL 12.3 Değişken A'nın (a_1, a_2) ve Değişken B'nin (b_1, b_2, b_3) bir fonksiyonu olarak tepki ortalaması



	b_1	b_2	b_3
a_1	19	19	20
a_2	10.6	15.8	18.2

Güven aralıkları (çubuklar) ± 1.44 birimdir.

Şekil 12.3, yukarıda değindiğimiz hipotetik deneyde yer alan altı ortalama etrafındaki güven aralıklarını gösteriyor. *GA*'nın incelenmesi tahminlerimizin kesinliği hakkında bize bilgi vermektedir. Ortalamalar etrafındaki aralıkların çakışıp çakışmadığına ve çakışıyorlarsa ne ölçüde çakıştıklarına bakarak aralık genişliğini ve *evren* ortalamalarının olası örüntülerini incelemek istiyoruz. Güven aralıklarını yorumlarken kullanılan pratik kurala göre, ortalamalar etrafındaki aralıklar çakışmıyorsa, SHAT kullanılarak test edildiğinde iki ortalamanın farklılaşma eğiliminde olduğunu hatırlayın (Bölüm 11'deki Kutu 11.5'e bakınız).

KARMA DESEN İÇİN İKİ FAKTÖRLÜ VARYANS ANALİZİ

Karma (mixed) desen için iki faktörlü varyans analizi, bir bağımsız değişken seçkisiz gruplar veya doğal gruplar desenini, ikinci bağımsız değişken ise tekrarlı ölçümler desenini temsil ettiği durumda uygundur. Birinci bağımsız değişken denekler arası faktör (burada A olarak sembolize edilmiştir) adını alır. İkinci bağımsız değişken ise denek içi faktör (burada B olarak sembolize edilmiştir) olarak adlandırılır. Karma desen için iki faktörlü analiz, tek faktörlü bağımsız gruplar deseni ve tek faktörlü tekrarlı ölçümler deseninin bir tür melezidir. Bu özel karmaşık desen, Bölüm 8'de, duygusal Stroop testinin kullanıldığı, Kaiser ve arkadaşlarına (2006) ait çalışmanın sonuçlarını rapor ettiğimizde tartışılmıştı.

Artık farkında olduğunuz gibi, ANOVA Özet Tablosu'na bakmadan önce ilgili özet istatistikleri gözden geçirmek ve verideki eğilimleri değerlendirmek önemlidir. Karma desen için iki faktörlü varyans analizine ait tipik bir bilgisayar çıktısının özeti aşağıda sunulmuştur. Bu analizin verilerini sağlayan deneyin ayrıntıları şu anda bizi ilgilendirmemektedir. Bazı bilgisayar yazılımlarının ilk önce gruplar arası analize, sonra da denek içi analize ait (etkileşimi de içeren) çıktıyı vererek,

Kaynak	*KT*	*sd*	*OK*	*F*	*p*
Denekler Arası					
Grup	0.225	1	0.225	1.718	0.226
Hata	1.049	8	0.131		
Denek İçi					
Sunum	15.149	2	7.574	58.640	0.000
Sunum × Grup	0.045	2	0.022	0.173	0.843
Hata	2.067	16	0.129		

karma desene ait çıktıyı ayırdığına dikkat edin. Tüm bilgiyi görebilmek için ekrandaki çıktı tablosunu kaydırmak zorunda olduğunuz dikkatinizi çekecektir.

Özet tablo iki bölüme ayrılmıştır. "Denekler arası" bölümü, grupların temel etkisi için *F*-oranını içermektedir. Tablonun bu kısmının şekli, bağımsız gruplar deseni için tek faktörlü analize benzemektedir. Bu bölümde listelenen hata, grup içi değişkenliktir. Grubun etkisine ait *F*-testi, istatistiksel olarak anlamlı değildir çünkü elde edilen .226 düzeyindeki olasılık, istatistiksel anlamlılığa ilişkin geleneksel .05 düzeyinden büyüktür. Özet tablonun ikinci bölümünün başlığı "Denek içi"dir. Bu bölüm denek içi değişken olan sunum sıklığının ("Sunum") temel etkisini ve sunum sıklığı ile grubun etkileşimini içermektedir. Genel olarak, denek içi değişkeni içeren herhangi bir etki (temel etki veya etkileşim) denek içi desende kullanılan artık hata terimi ile test edilmelidir. Buradaki etkileşime ait *F*-testi değeri 1'den küçüktür istatistiksel olarak anlamlı değildir. Bununla birlikte, sunum sıklığının temel etkisi istatistiksel olarak anlamlı bir *F* değeridir (Tek faktörlü denek içi desenin analizinde olduğu gibi, bilgisayar çıktınız burada belirttiklerimizin dışında ek bilgiler verebilir.).

Karma desen için iki faktörlü analizin sonuçlarının yorumlanması, herhangi bir karmaşık desendeki mantığı izler. Bununla birlikte, bir karma deseni analiz ederken, özet tabloda listelenen analizlerin ötesindeki analizler için (yani basit temel etkiler, ikili ortalama karşılaştırmaları) uygun hata teriminin kullanılmasına özen gösterilmelidir. Örneğin, anlamlı bir etkileşim elde edildiğinde basit temel etkilerin, her basit etki, ikinci bağımsız değişkenin o düzeyindeki tek faktörlü ANOVA gibi düşünülerek analiz edilmesi önerilmektedir. Örneğin, deneyimizde grup ve sunum sıklığı arasında anlamlı bir etkileşim elde etmişsek, işlem yapılan grup için basit temel etki, sadece o grup için tekrarlı ölçümler ANOVA yapılmasını içerecektir (bu tür karşılaştırmalara ait daha fazla bilgi için bk. Keppel, 1991).

Karma desendeki etki büyüklüğü tahminleri de sıklıkla eta kareyi, yani bağımsız değişken tarafından açıklanan değişkenliğin oranına dair tahmini kullanır. Görmüş olduğunuz gibi eta kare, söz konusu etki için, *KT* etkisinin, *KT* etkisi ve *KT* hatanın toplamına bölümü olarak tanımlanır.

KARMAŞIK DESEN İÇİN SONUÇLARIN RAPOR EDİLMESİ

Karmaşık bir desenin sonuçlarının rapor edilmesi, tek faktörlü ANOVA raporunun genel yapısını izler ama etkileşim varsa, bu etkileşimin niteliğine özellikle dikkat edilir. Aşağıdakiler, bir karmaşık desenin sonuçlarına ait raporun önemli ögeleridir:

- değişkenlerin açıklanması ve her bir değişkenin düzeylerinin (koşulların) belirlenmesi;
- desen matrisindeki hücrelerin özet istatistiklerinin, eğer uygunsa grup ortalamalarına ait güven aralıklarıyla birlikte, metin içinde, tabloda veya şekilde yer alması;
- temel etkiler ve etkileşimlere ait tam olasılıklarla birlikte, F-testlerinin rapor edilmesi;
- her etkinin etki büyüklüğü ölçüleri;
- anlamlı olmayan etkilere ait gücün bildirilmesi;
- etkileşim istatistiksel olarak anlamlıysa, basit temel etki analizleri;
- istatistiksel olarak anlamlı olan etkileşimin (eğer varsa), okuyucuyu, bağımsız değişkenlerin düzeyleri boyunca hücre ortalamaları arasındaki farklara yönlendirerek sözel ifadesinin yapılması
- okuyucuyu, istatistiksel olarak anlamlı olan temel etkinin (eğer varsa), bağımsız değişkenlerin düzeyleri boyunca birleştirilen hücre ortalamaları arasındaki farklara yönlendirerek, sözel olarak ifade edilmesi –temel etkiye katkı yapan, ortalamalar arasındaki sistematik değişkenliğin kaynaklarını açıklamak üzere, uygun olduğunda, ikili ortalama karşılaştırmalarının yapılması;
- bu analizin bulgularından hareketle, okuyucunun ulaşmasını istediğiniz sonucun ifadesi.

Sonuçlar bölümünün APA stiline göre nasıl yazılacağına ait ek ipuçları Bölüm 13'tedir.

ÖZET

Sıfır hipotezi anlamlılık testini (SHAT) temel alan istatistiksel testler psikolojide genellikle onaylayıcı veri analizi için kullanılır. SHAT, bir deneyde bağımsız değişkenlerin yarattığı farkların, tamamen hata değişkenliğinin (şans eseri) doğurabileceği farklardan büyük olup olmadığını belirlemek için kullanılır. Sıfır hipotezi, bağımsız değişkenin etkisinin olmadığını öne süren hipotezdir. İstatistiksel olarak anlamlı bir sonuç, sıfır hipotezinin doğru olma olasılığının küçük olduğu bir sonuçtur. SHAT yapılırken iki tip hata ortaya çıkabilir. Araştırmacı, sıfır hipotezini doğruyken reddettiğinde Tip I hata ortaya çıkar. Tip I hatanın olasılığı alfaya, yani anlamlılık düzeyine eşittir ve genellikle .05'tir. Tip II hata ise, yanlış bir sıfır hipotezi reddedilmediği zaman ortaya çıkar. Tip II hatalar, bir çalışma, sıfır hipotezini doğru bir şekilde reddedebilmek için yeterli güce sahip olmadığında ortaya çıkabilir. Gücü artırmak için araştırmacıların kullandığı başlıca yol, örneklemi büyütmektir. Araştırmacılar güç tablolarını kullanarak, çalışma yapılmadan önce, yanlış bir sıfır hipotezini reddedebilmek için gerekli olan gücü ve çalışma tamamlandıktan sonra, bulunan etkiyi saptama olasılığını tahmin edebilirler. Bir istatistiksel testin sonucuna ilişkin tam olasılık rapor edilmelidir.

İki ortalamanın karşılaştırılması için uygun olan istatistiksel test, *t*-testidir. İki ortalama arasındaki fark test edildiğinde, Cohen *d* gibi bir etki büyüklüğü ölçüsü de rapor edilmelidir. APA *Yayın Kılavuzu*, güven aralıklarının SHAT sonuçlarıyla birlikte rapor edilmesini özellikle önermektedir. SHAT'nin sonuçları rapor edilirken, istatistiksel anlamlılığın (veya çakışmayan güven aralıklarının), bilimsel ya da pratik anlamlılıkla aynı şey olmadığını unutmamak önemlidir. Dahası, ne SHAT ve güven aralıkları ne de etki büyüklükleri bize bir çalışmanın yönteminin sağlamlığına ilişkin bir şey söylemez. Başka bir deyişle, bu ölçümlerin hiç birisi tek başına alternatif bir hipotezin (bağımsız değişkenin etkisinin olduğu hipotezinin) doğru olduğunu öne sürmek için kullanılamaz. Ancak analiz sonuçlarının elde edilmesinde kullanılmış olan yöntemi dikkatle inceledikten sonra davranışı neyin etkilemiş olduğuna dair bir iddia ortaya atabiliriz.

Varyans analizi (ANOVA), üç veya daha fazla ortalamayı karşılaştırmaya uygun bir istatistiksel testtir. ANOVA'nın temelindeki mantık, verideki hem hata değişkenliğini, hem de sistematik değişkenliğin kaynaklarını belirlemeyi temel alır. Hata değişkenliği ve sistematik değişkenliğin (varsa), tek başına hata değişkenliğine bölünmesini temsil eden bir *F*-testi yapılandırılmaktadır. Genel analizin sonuçları, genel *F*-testi olarak adlandırılır ve ANOVA Özet Tablosu'nda rapor edilir. Büyük bir *F*-oranı, bağımsız değişkenin etkisinin olduğu yönünde kanıt sağlar. Tek faktörlü bağımsız gruplar deseni için etki büyüklüğü ölçüleri arasında Cohen *f* ve eta kare (η^2) yer alır. İkili ortalama karşılaştırmaları, anlamlı bir *F*-testine katkıda bulunan sistematik değişkenliğin kaynaklarını daha açık şekilde belirlemek amacıyla, genel *F*-testinin sonuçları alındıktan sonra yapılır. Güven aralıkları, çoklu grup çalışmasından elde edilen verilerle yapılan ANOVA'yı tamamlamak için kullanılabilir ve SHAT sonuçları özetlendiğinde rapor edilmelidir.

İki faktörlü ANOVA, bir araştırmacı karmaşık bir desende aynı anda iki veya daha fazla bağımsız değişkenin davranış üzerindeki etkisini araştırdığı zaman uygundur. Bağımsız değişkenlerden birisi gruplar arası bir değişkeni (seçkisiz veya doğal gruplar), diğeri ise tekrarlı ölçümler denek içi değişkeni temsil ettiği zaman, karma desen ortaya çıkar. Değişkenlere ilişkin hem temel etkileri hem de etkileşimi değerlendirmek için genel *F*-testi yapılır. İstatistiksel olarak anlamlı bir etkileşim bulunduğunda, etkileşimin kaynağı, temel etki analizlerinin yapılmasıyla araştırılabilir. Basit temel etki, bir bağımsız değişkenin ikinci bağımsız değişkenin sadece bir düzeyindeki etkisidir. Güven aralıkları da bir karmaşık desendeki bağımsız değişkenin etkisini anlamada yardımcı olarak kullanılabilir. Karmaşık desende sık kullanılan bir etki büyüklüğü ölçüsü eta karedir.

TEMEL KAVRAMLAR

tekrarlı ölçümler (denek içi)
anlamlılık düzeyi 381
sıfır hipotezi (H_0) 381
Tip I hata 383
Tip II hata 383
güç 384
bağımsız gruplar t-testi 387
t-testi 387
ANOVA 393
tek faktörlü bağımsız gruplar deseni 393
F-testi 394
genel F-testi 394
Cohen f 399
eta kare (η^2) 398
iki ortalamanın karşılaştırılması 401

GÖZDEN GEÇİRME SORULARI

1. Bir istatistiksel testin sonuçlarının "istatistiksel olarak anlamlı" olduğunu söylemek ne anlama gelir?
2. SHAT yapılırken oluşan Tip I ve Tip II hatalar arasındaki farkı belirtin.
3. Bir istatistiksel testin gücünü belirleyen üç faktör hangisidir? Araştırmacıların gücü kontrol etmek amacıyla kullanabileceği başlıca faktör hangisidir?
4. Tekrarlı ölçümler deseni neden seçkisiz gruplar desenine göre daha duyarlı olma eğilimindedir?
5. Etki büyüklüğünü kullanmanın bir avantajını betimleyin ve bir sonuç istatistiksel olarak anlamlı değilse, güç analizinin nasıl kullanılabileceğini açıklayın.
6. İstatistiksel olarak anlamlı olan bir sonuç neden, hem bilimsel, hem de pratik olarak anlamlı olmayabilir?
7. *F*-testinin mantığını kısaca özetleyin.
8. Genel *F*-testinden elde ettiğiniz bilgi ile ikili ortalama karşılaştırmalarından elde ettiğiniz bilgi arasındaki fark nedir?
9. Tekrarlı ölçümler ANOVA ile bağımsız gruplar ANOVA'nın farklılaştığı esas nokta nedir?
10. Basit temel etki, genel temel etkiden nasıl ayrılır?

DÜŞÜNDÜRÜCÜ SORULAR

1 Bir araştırmacı, küçük çocuklara okuma öğretmek için iki farklı yöntemi karşılaştıran bir deney yapmıştır. Eski bir yöntem, yeni bir yöntemle karşılaştırılmıştır ve yeni yönteme ait ortalama performansın, eski yöntemin performansından yüksek olduğu bulunmuştur. Sonuçlar, $t(120) = 2.10$, $p = .04$ ($d = .34$) şeklinde rapor edilmiştir

- **A** Sonuç istatistiksel olarak anlamlı mıdır?
- **B** Bu çalışmada kaç katılımcı yer almıştır?
- **C** Etki büyüklüğü ölçüsü olan d'ye göre, bu çalışmada bulunan etkinin büyüklüğü hakkında ne söyleyebiliriz?
- **D** Araştırmacı bu sonucu temel alarak, küçük çocuklara okuma öğretmede yeni yöntemin açıkça pratik anlamlılığa sahip olduğunu ve bir an önce uygulanmasını ifade etmiştir. Böyle bir ifadeye tepkiniz ne olur?
- **E** Güven aralıklarının yapılandırılması, bu sonuçları anlamamıza nasıl bir katkıda bulunur?

2 Bir sosyal psikolog, üç propaganda mesajının üniversite öğrencilerinin terörizme karşı savaşla ilgili tutumları üzerindeki etkisini karşılaştırmıştır. Doksan (N = 90) öğrenci üç farklı iletişim koşuluna, seçkisiz olarak eşit sayıda atanmıştır. Öğrenciler propagandaya maruz kaldıktan sonra savaşla ilgili tutumlarını değerlendirmek için bir kâğıt-kalem tutum ölçümü kullanılmıştır.

Kaynak	Kareler Toplamı	*sd*	Ortalama Kare	*F*	*p*
İletişim	180.10	2	90.05	17.87	0.000
Hata	438.50	87	5.04		

Üç mesajın, öğrencilerin tutumları üzerindeki etkisini belirlemek için ANOVA yapılmıştır. ANOVA Özet Tablosu şöyledir:

- **A** Sonuç istatistiksel olarak anlamlı mıdır? Neden anlamlıdır veya değildir?
- **B** Bu sonuçlardan hareketle, hangi etki büyüklüğü kolayca hesaplanabilir? Bu ölçümün değeri kaçtır?
- **C** İkili ortalama karşılaştırmaları yapmak bu sonuçların yorumlanmasına nasıl katkı sağlayabilir?
- **D** Grup ortalamaları verilmemiş olsa da bu verilerden hareketle, birleştirilmiş değişkenlik tahminini temel alarak, ortalamalar için güven aralığının genişliğini hesaplamak mümkündür. Bu çalışmada ortalamalar için güven aralığının genişliği nedir?

3 Bir gelişim psikoloğu 4, 6 ve 8'inci sınıftaki çocuklara iki tür eleştirel düşünme testi vermiştir. Her sınıfta test edilen 28 çocuk vardır;

(*devam ediyor*)

14'ü testin bir formunu (A veya B) almıştır. Bağımlı değişken, testteki doğru yüzdesidir. Her

Test	4'üncü	6'ncı	8'inci
Form A	38.14	63.64	80.21
Form B	52.29	68.64	80.93

sınıftaki çocukların ortalama doğru yüzdesi aşağıdadır:

Kaynak	Kareler Toplamı	*sd*	Ortalama Kare	*F*	*p*
Not	17698.95	2	8849.48	96.72	.000
Test	920.05	1	920.05	10.06	.002
Not × Test	658.67	2	329.33	3.60	.032
Hata	7136.29	78	91.49		

Bu deney için ANOVA Özet Tablosu aşağıdadır::

- **A** Bu deneydeki ortalamaları gösteren bir grafik çizin. Grafiğe ilişkin incelemenizi temel alarak, değişkenler arasındaki istatistiksel olarak anlamlı bir etkileşimden şüphelenir miydiniz? Neden şüphelendiğinizi veya şüphelenmediğinizi açıklayın.
- **B** Hangi etkiler istatistiksel olarak anlamlıdır? İstatistiksel olarak anlamlı olan her bir etkiyi sözel olarak açıklayın.
- **C** Sınıf derecesi ve testin temel etkisi için eta kare değerleri nedir?
- **D** Etkileşimin kaynağını belirlemek için hangi ek analizleri yapabilirsiniz?
- **E** Her bir sınıf için testin basit temel etkisi nedir?
- **F** Deneydeki altı ortalama için güven aralıklarını hesaplayın ve bu güven aralıklarını, bu sonuçlara ilişkin grafiğinizde ortalamaların etrafında çizin.

İleri Alıştırmanın Cevapları

İfade 1 ve 5 doğru; 2, 3 ve 4 yanlış.

Düşündürücü Soruların Cevapları

A. Evet. Bu sonuç için elde edilen olasılık, sıfır hipotezinin doğru olduğu varsayımı altında, anlamlılığa ilişkin geleneksel düzey olan .05'ten küçüktür.

B. Rapor edilen serbestlik derecesi (*sd*) 120'dir. Bağımsız gruplar *t*-testi için, $sd = n_1 + n_2 - 2$'dir. Bu nedenle 122 katılımcının olması gerekir.

C. Cohen'in yönergesi, .20 düzeyindeki etki büyüklüğünün küçük bir etki, .50'nin orta düzeyde veya ortalama bir etki, ve .80'in ise büyük bir etki olduğunu belirtiyor. .34 düzeyindeki bir etki büyüklüğü ise küçük ve orta düzeydeki etki arasındadır.

D. SHAT'nin sonuçları, pratik anlamlılık hakkında doğrudan bir şey söylemez. Eğer yeni yöntem çok daha pahalıysa, uygulamak için çok zaman alıyorsa veya hemen temin edilemeyecek kaynakları (örneğin, yeni okuma materyalleri) gerektiriyorsa, bu bulgunun pratik anlamlılığının (en azından kısa vadede) küçük olması olasıdır. Etki büyüklüğü de geniş olmadığı için bu durum özellikle böyle olabilir. Ayrıca $p = .04$ olması da .05 düzeyinde istatistiksel olarak anlamlı olan bu sonucun tekrar edilebilme olasılığının o kadar yüksek olmadığını işaret etmektedir. Son olarak, çalışmanın eksiksiz veya karıştırıcı değişkenlerden ve deneyci hatalarından arınmış olup olmadığını belirlemek için, çalışmanın yöntemini dikkatle incelemek isteriz.

E. İki evren ortalaması için bir güven aralığının yapılandırılması, bu yöntemler arasındaki farkın büyüklüğüne ilişkin bir kanıt sunacaktır ve iki evren ortalaması arasındaki farka ilişkin tahminin kesinliğini gösterecektir (aralığın genişliğinin incelenmesini temel alarak).

BÖLÜM ON ÜÇ

Psikolojide İletişim

BÖLÜMÜN ANA HATLARI

GİRİŞ
İNTERNET VE ARAŞTIRMA
ETKİLİ YAZIM İÇİN ÖNERİLER
ARAŞTIRMA RAPORUNUN YAPISI
Başlık Sayfası
Özet
Giriş
Yöntem
Sonuçlar
Tartışma
Kaynaklar
Dipnotlar
Ekler
SÖZEL SUNUMLAR
ARAŞTIRMA ÖNERİLERİ

GİRİŞ

Bilimsel araştırma herkese açık olarak gerçekleştirilen bir etkinliktir. Zekice bir hipotez, zarif bir araştırma deseni, kılı kırk yaran veri toplama işlemleri, güvenilir sonuçlar ve bulguların ustaca yapılmış kuramsal yorumu, herkese açık hâle getirilmedikçe, bilim camiası için kullanışlı değildir. Bir yazarın güçlü bir biçimde vurgulayarak önerdiği gibi "Bilimsel araştırma, sonuçları, tercihen yüksek standartlı bir hakemli dergide, sancılı yayın sürecinden geçene kadar, sadece bir oyundur. Yayın, bilimin vazgeçilmez bir parçasıdır" (Bartholomew, 1982, s. 233).

Bartholomew, "hakemli" dergi için bir tercih belirtmişti çünkü hakemli dergiler *bağımsız değerlendirme* sürecinden geçmektedir. Gönderilen makale taslakları belirli bir araştırma alanında uzman olan başka araştırmacılar tarafından ("bağımsız araştırmacılar") incelenmektedir. Bu araştırmacılar araştırmanın yöntemsel açıdan uygun olup olmadığına ve psikoloji disiplinine önemli bir katkı sağlayıp sağlamadığına karar vermektedirler. Daha sonra bu kararlar derginin editörü olarak görev yapan daha kıdemli araştırmacıya gönderilir. Hangi yazıların yayımlanmaya hak kazandığına karar vermek editörün işidir. Bağımsız değerlendirme, psikoloji alanında yayımlanmış bir araştırmanın kalite kontrolü için başlıca yöntemdir.

Araştırmacıların bulgularını yayımlayabildiği çok sayıda psikoloji dergisi mevcuttur. *Psychological Science, Memory & Cognition, Child Development, Journal of Personality and Social Psychology, Psychological Science in the Public Interest* ve *Journal of Clinical and Consulting Psychology* bunlardan sadece bir kaçıdır. Editörün, hangi makale taslaklarının yayımlanacağına dair kararı, editör ve bağımsız araştırmacılar tarafından değerlendirilen; (a) araştırmanın kalitesine ve (b) araştırmanın yazılı taslaktaki sunumunun etkililiğine bağlıdır. Bu nedenle hem içerik, hem de biçim önemlidir. Editörler, açık bir şekilde betimlenmiş en iyi araştırmaları aramakta ve kabul için katı standartlar koymaktadırlar. *Genellikle, çok sayıdaki APA dergisi arasından, ikiden fazla dergiye gönderilen her dört makale taslağının yaklaşık* sadece biri yayın için kabul edilmektedir (örneğin, Amerikan Psikoloji Birliği, 2012).

Bir dergi editörü, bir makale taslağını biçim ve içerik açısından değerlendirmeye ek olarak öncelikle, gönderilen taslağın o dergi için uygun olup olmadığına karar verecektir. Örneğin, hayvanlarla yapılan deneysel bellek çalışmaları çoğunlukla, çocuk gelişimi ile ilgili araştırmalara ağırlık veren dergilerde yayımlanmamaktadır. APA ve APS'nin desteklediği dergilerden başka, yayın için birçok kaynak bulunmaktadır. Bununla birlikte, tam olarak ne var ne yok anlamaya başlamak için şu başlıca kuruluşların yayımladığı dergilerin açıklamalarını gözden geçirmek isteyebilirsiniz: www.apa.org/pubs/journals/ ve www.psychologicalscience.org/journals/.

Editör incelemesi ve yayın süreci uzun zaman alabilir. Bir makalenin gönderilmesi ile en sonunda dergide yayımlanması arasında bir yıla kadar (ve bazen daha da uzun) zaman geçebilir. Makalenin kabul edilip edilmeyeceğine dair karar verilmeden önce makale taslağının incelenmesi birkaç ay sürebilir. Ayrıca yayın süreci için, makalenin kabul edilmesi ile dergide fiilen yayımlanması arasında bir kaç ay daha gereklidir. Dergi makaleleri, gittikçe artan bir oranda, basılmadan önce çevrim içi olarak (internette) yayımlanmaktadır. Araştırma bulgularının zamanında rapor edilmesini sağlayabilmek için Amerikan Psikoloji Birliği, Psikoloji Bilimi Birliği, Psikoloji Topluluğu, Çocuk Gelişimi Araştırma Topluluğu gibi profesyonel topluluklar ve Doğu, Orta Batı, Güney Doğu ve Batı'daki psikoloji birlikleri

gibi yerel topluluklar, araştırmacıların, son çalışmalarını betimleyen kısa sözlü sunumlar yaptıkları veya "posterler" sundukları konferansları desteklemektedir. Bu türden konferanslar, aynı araştırma soruları ile ilgilenen araştırmacılar arasında tartışma ve fikir alışverişinin zamanında yapılması için fırsat sunmaktadır. "Baskıda" olan (yani yayın sürecinin tamamlanmasını bekleyen) araştırma tartışılabilir, böylece konferans katılımcılarına önemli fakat yayımlanmak üzere olan, henüz yayımlanmamış araştırma bulgularının bir ön sunumu yapılır.

Araştırmacılar genellikle araştırmalarını yürütmek için bir devlet kuruluşundan veya özel bir kuruluştan, ödenek biçiminde finansal destek almaktadırlar. Ödenekler, araştırma önerilerinin rekabete dayalı bir şekilde incelenmesi temelinde, ödül olarak verilmektedir. Ayrıca araştırma önerileri genellikle yüksek lisans veya doktora tezi hazırlarken lisansüstü öğrencileri için zorunluluktur. Akademik bir kurul yüksek lisans veya doktora tezi başlamadan önce öneriyi incelemektedir. Lisans öğrencilerinin de psikolojide araştırma yöntemleri veya laboratuvar dersinin bir parçası olarak sıklıkla araştırma önerisi hazırlamaları gerekmektedir. Son olarak, her düzeydeki araştırmacı, araştırma önerilerinin, etik kurullar tarafından, önerilen araştırmanın etik niteliği açısından bir kurumda değerlendirilme zorunluluğu olduğunu fark edecektir (bk. Bölüm 3). Araştırma önerilerinin, tamamlanmış bir çalışmanın sonuçlarını rapor eden dergi makalesinden biraz farklı bir stil ve biçime sahip olması gerekmektedir. Bu bölümde daha sonra, araştırma önerisi hazırlamakla ilgili önerilerde bulunacağız.

Makale Taslağı Hazırlamaya İlişkin İpuçları Psikolojide bilimsel yazım için başlıca kaynak, *Amerikan Psikoloji Birliği Yayın Kılavuzunun* (APA, 2010b) altıncı baskısıdır. Dergi editörleri ve yazarlar, psikolojideki birçok farklı derginin tutarlı bir stile sahip olduğundan emin olmak için bu kılavuzu kullanmaktadırlar. Kılavuz, bir psikoloji dergisinde yayımlanmak üzere hazırlanan makale taslağının stili ve biçimiyle ilgili sorulabilecek hemen hemen her soru için paha biçilemez bir kaynaktır. Uygun içerik ve makale taslağının organizasyonu; düşüncelerin ifadesi ve dildeki yanlılığın azaltılması; sonuçların tablolarda ve şekillerde gösterilmesi; elektronik medyaya referans verme dâhil, kaynak listesinin biçimi ve makale taslaklarının elektronik olarak gönderilmesini de içeren, makale taslağı kabulü ve üretilmesi ile ilgili ilkeler hakkında bilgi içerir. Kılavuz ayrıca, bilimsel yazım ile ilgili etik meseleleri tartışmaktadır (Bölüm 3'te konu ile ilgili tartışmamıza bakınız). Bununla birlikte APA, ne editöryal stilin, ne de yayın teknolojisinin durağan olmadığını da kabul etmektedir. APA kılavuzuna göre makale taslağı hazırlamayı amaçlayan herhangi biri, ayrıca *Yayın Kılavuzu* ile ilgili güncellemeleri ve APA stili ile APA ilkeleri ve prosedürlerindeki son değişiklikleri sunan APA'nın web sayfasına da başvurmalıdır: www.apastyle.org.

APA'nın web sayfası, örnek makale taslağı gösterimi ve sık sorulan sorulara verilen yanıtları da içeren temel APA stili ile ilgili ücretsiz bir eğitim materyali niteliğindedir.

Peki, dergi makaleleri, sözel sunumlar ve araştırma önerileri ne işinize yarayacak? Eğer psikolojide lisansüstü bir programa devam ediyorsanız, muhtemelen bu üç bilimsel iletişim türünün hepsini kullanarak araştırmanızı betimlemek zorunda kalacaksınız. Psikolojide profesyonel bir kariyer yapmanın peşinde olmasanız bile,

iyi bir yazılı ve sözlü araştırma raporunun ilkeleri çok çeşitli meslek alanlarına uygulanabilir. Örneğin, bölüm yöneticiniz için son satışlara ait sonuçları betimleyen bir rapor, kısa bir dergi makalesiyle neredeyse aynı biçime sahip olabilir. Daha da önemlisi, araştırma yöntemleri dersinizde araştırma önerisi hazırlamak ve araştırma raporu yazmak veya teslim etmek zorunda olabilirsiniz. Bu bölüm, bu türden işleri gerektiği gibi yapabilmeniz için size yardımcı olacaktır.

Bu bölüm, öncelikle makale taslağı hazırlamada ilk adımı atmanız için size yardımcı olmayı amaçlamaktadır ve *Amerikan Psikoloji Birliği Yayın Kılavuzu'nun* (2010b) yerine geçemez. Aşağıdaki kısımlar *Kılavuzun*, yazarlar ve yayıncı (McGraw-Hill) tarafından yapılmış bir yorumudur ve en güncel, kesin APA stili için, *APA Kılavuzunun* son baskısına ve APA'nın web sayfasına başvurmanızı öneriyoruz.

İNTERNET VE ARAŞTIRMA

İnternet hâlihazırda, araştırma gerçekleştirmek için vazgeçilmez bir araç hâline gelmiş durumdadır fakat aynı zamanda bilim camiasının üyeleri arasında iletişim için de esastır. Birçok araştırmacı için *e-posta*, çalışma arkadaşları, dergi editörleri, araştırma ortakları, mali destek veren temsilciliklerin yöneticileri ve diğer profesyoneller ile iletişimin başlıca yoludur. Az önce okuduğunuz makale ile ilgili bir sorunuz mu var? Bir e-posta mesajı göndererek yazara sorun. E-posta göndermek basit, etkili ve uygundur. Örneğin kitabınızın yazarlarına, ZechResearchMethods@gmail.com adresine e-posta mesajı gönderilerek ulaşılabilir.

Ayrıca web'de bu kitap için tahsis edilmiş, öğrenci materyalleri (örneğin, alıştırma testleri) ve yazarlar, baskılar arasındaki değişiklikler, psikoloji araştırmaları yapmak için ek kaynaklar ve güncel baskıdaki hatalar ya da eksiklikler ile ilgili bilgiler, yayıncı adresi, sipariş bilgileri vs. için ulaşılabilecek bir sayfa mevcuttur. Sayfamızı ziyaret edin: www.mhhe.com/shaughnessy10.

İnternet ayrıca öğrencilere ve profesyonel psikologlara tartışma grupları, veri tabanları ve elektronik dergiler dâhil başka birçok önemli şekilde hizmet etmektedir.

"Liste sunucuları" olarak adlandırılan *tartışma grupları*, ilgilenen kişilerin, psikolojiye dair ortak ilgi alanını paylaştıkları meseleleri tartışmalarını sağlamaktadır. Grup, devam etmekte olan bir tartışmaya katkıda bulunmak isteyen "abonelerin" bir "listesinden" oluşmaktadır. Liste üyelerine, bir abone tarafından gönderilen herhangi bir mesaj anında "sunulmaktadır". İnternette, bağımlılık, din ve kadın çalışmaları gibi, dünya çapında geniş çeşitlilikteki konuları tartışan araştırmacıları birbirine bağlayan yüzlerde liste sunucusu bulunmaktadır. Bazı liste sunucuları tartışmaya katılmak isteyen, sadece pasif olarak katılmak isteyenler de dâhil ("lurk-pusuda bekleyen"), herkese açıktır. Diğer liste sunucuları ise sadece belirli yeterliliğe sahip bireylere açıktır (örneğin, belirli bir APA biriminin üyeleri). APA ve APS ayrıca öğrenciler için www.apa.org/apags/ ve www.psychologicalscience.org/apssc/ adreslerinden ulaşılabilen tartışma gruplarını da desteklemektedir.

İnternetteki *veri tabanları* şunlardır: İnternette depolanan ve elektronik olarak ulaşılabilen elektronik veri dosyaları. Sadece birkaç tanesini belirtmek gerekirse, örneğin tıp, alkolizm, kamuoyu yoklamaları ile ilgili veri tabanları mevcuttur. Veri tabanları, özellikle arşiv araştırması (bk. Bölüm 4) ve zaman dizileri analizi (Bölüm 10) yaparken kullanışlıdır. Yüzlerce değişkenin ve çok sayıda katılımcının ulaşılabilir olduğu büyük çaplı veri tabanları, psikolojideki araştırma sorularını yanıtlamayı

amaçlayan birçok araştırmacı için önemli hâle gelmiştir. Veri tabanlarına elektronik olarak erişim, araştırmacıları, zaten ulaşılabilecek olan verileri toplamak için gerekli olan masraf ve zamandan kurtarmaktadır; bu sayede araştırmacıların ve katılımcıların lüzumsuz yere çaba harcamalarının önüne geçilmektedir.

Elektronik dergiler yaygındır ve makale taslaklarının elektronik olarak gönderilmesi, bugün dergiler ve konferanslar için bir kural hâline gelmiştir. E-posta ve İnternet erişiminin yaygın ulaşılabilirliği, değerlendirme sürecini hızlandırmıştır; şöyle ki, makale taslağının gönderilmesi, hakem değerlendirmeleri ve editörün yazarlara olan geri bildirimi İnternet aracılığıyla yapılabilmektedir. Buna ek olarak, bazı dergiler sadece elektronik biçimde sunulmaktadır. Aboneler, makaleleri elektronik posta kutularından almaktadır ve okuyucular makaleler ile ilgili yorumlarını elektronik olarak gönderebilmektedir. *Current Research in Social Psychology* ve *Prevention and Treatment*, elektronik dergilere örnektir. Araştırmalarını prestijli (etki katsayısı yüksek) dergilerde yayımlamayı amaçlayan yazarlar, makale taslaklarını elektronik veya basılı dergilere göndermekten bağımsız olarak, hakem değerlendirmesinden geçmesini beklemelidirler.

ETKİLİ YAZIM İÇİN ÖNERİLER

İyi yazmayı öğrenmek, yüzmeyi, araba kullanmayı veya piyano çalmayı öğrenmeye benzer. Gelişmenin, sadece aktivitenin nasıl yapılacağını okumaya bağlı olması olası değildir. Uzman önerilerine önem vermek, bir kimsenin iyi bir başlangıç yapmasına yardımcı olabilmektedir. Bu nedenle iyi yazabilmenin bir anahtarı da yazma "koçlarından" –öğretmenlerden, arkadaşlardan, editörlerden ve hatta kendinizden– eleştirel geri bildirim almaktır.

İyi yazım, iyi araç sürmede olduğu gibi, en iyi şekilde savunmacı bir tarzla gerçekleştirilebilir. Şunu varsayın: Yanlış anlaşılma olasılığı olan her şey, yanlış anlaşılacaktır! Bu tür yazım kazalarından kaçınmak için, yazmaya başlamadan önce aşağıdaki ipuçlarını dikkate alın.

- **OKUYUCUNUZU İYİ TANIYIN:** Eğer okuyucularınızın, gerçekte olduğundan daha fazla şey bildiğini varsayarsanız, okuyucularınızın zihninin karışmasına neden olursunuz. Eğer okuyucularınızı hafife alırsanız da, onları gereksiz ayrıntılarla sıkma riski ile karşı karşıya kalırsınız. Her iki risk de, yazdığınız şeyin okunmama olasılığını artıracaktır. Eğer hata yapmanız gerekiyorsa, okuyucularınızı hafife almak daha iyi bir seçenektir.
- **AMACINIZI BELİRLEYİN:** Dergi makaleleri, genel bir kategori olan bilgilendirici yazım kategorisine dâhildir. Webster Sözlüğü, bilgilendirmeyi "bilgi iletmek veya anlaşılması güç olan bir şeyi izah etmek için tasarlanmış ifade" olarak tanımlamaktadır. Dergi makalesinin temel amaçları, betimlemek ve ikna etmektir. İlk olarak ne yaptığınızı ve ne bulduğunuzu betimlemek, ikinci olarak da okuyucuyu, söz konusu sonuçlara getirdiğiniz yorumun uygun olduğu konusunda ikna etmeyi istersiniz.
- **AÇIK YAZIN:** İyi bir bilgilendirici yazımın esası, düşüncenin ve ifadenin açıklığıdır. Tanınmış bir editörün vurguladığı gibi, "Yazarken önemli olan şey itinadır (Cronbach, 1992, s. 391). Fikirlerinizin pürüzsüz ve mantıklı bir biçimde akışını sağlamak için cümleler üzerinden tekrar ve tekrar geçmeniz gerekecektir. *Yayın Kılavuzunda* belirtildiği gibi (s. 65), "Bilimsel rapor yazmanın başlıca amacı, açık bir iletişimdir."

- **KISA VE ÖZ OLUN:** Sadece söylenmesi gerekeni söylerseniz, ifade biçiminizde ekonomik davranmış olursunuz. Kısa sözcükler ve kısa cümlelerin anlaşılması okuyucular için daha kolaydır. Laf kalabalığını gidermenin en iyi yolu, kendi yazınızı her taslakta düzenlemeniz ve başkalarından yazınızın taslaklarını düzenlemesini istemenizdir.
- **KESİN OLUN:** Dilde kesinlik, söylemek istediğiniz şey için doğru sözcüğü seçmek anlamına gelmektedir. Bu durum, tam olarak istediğiniz anlamı veren sözcükleri seçmeyi gerektirir. Örneğin bilimsel psikolojide *inanç* ile *tutum*, veyahut *duyumlar* ile *hisler* aynı şey değildir.
- **DİL BİLGİSİ KURALLARINA UYUN:** Dilbilgisi kurallarına bağlılık, iyi bir yazım için kesinlikle gereklidir çünkü bunu yapamamak, okuyucunun ilgisini dağıtır ve anlam karmaşasına neden olabilir. Ayrıca sizde, yani yazarda, hoş olmayan bir görüntüye neden olur ve sonuçta okuyucunun gözünde güvenilirliğinizi (ve iddianızı) zayıflatır.
- **ADALETLİ YAZIN:** Yazar olarak ayrıca, bireyleri adil bir şekilde ve yansız olarak kabul edecek şekilde sözcükleri seçmeye ve yorum yapmaya gayret etmelisiniz. Amerikan Psikologlar Birliği, yazarların kullandığı dildeki yanlılık ile ilgili ilkesini özetlemiştir (*Yayın Kılavuzu*, 2010b, ss. 70–77):

  Bilimsel yazım, üzerinde çalışılan grup veya grupların örtük veya konu dışı değerlendirmelerinden arınmış olmalıdır. Bir kuruluş olarak APA kendini bilime ve bireylerin ve grupların adil olarak muamele görmesine adamıştır ve bu ilke, APA yayınları için yazan yazarların, yazılarında, kişiler hakkındaki, sürekli olarak alçaltıcı tutumlardan ve yanlı varsayımlardan kaçınmalarını gerektirir. Cinsiyet, cinsel yönelim, ırksal veya etnik grup, engellilik ya da yaş ile ilgili yanlılık anlamına gelebilecek yorumlar kabul edilemez. (ss. 70–71)

Yayın Kılavuzu (2010b, ss. 70–77), yansız bir iletişim yapmanıza yardımcı olacak önemli bilgiler sunmaktadır. Aşağıda sadece *Kılavuzda* yer alan önerileri temel alan kısa bir giriş yer almaktadır (ayrıca bk. www.apastyle.org):

(a) Bireyleri uygun düzeyde bir özgüllükte betimleyin. Örneğin, insan yetişkinlerden bahsederken kullanılan *erkekler* ve *kadınlar* ifadeleri, genel bir terim olan *adam (man)* teriminden daha doğrudur. "Çinli Amerikalı" veya "Meksikalı Amerikalı" ifadeleri, araştırma katılımcıları için, Asya kökenli Amerikalı veya İspanyol kökenli Amerikalı terimlerinden daha özgül bir gönderme olacaktır.

(b) Bireylerin ırksal veya etnik kimlikleri ile ilgili terimleri kullanırken olduğu gibi, kişilere atıfta bulunurken de etiketlere karşı duyarlı olun. Bu kılavuzu takip etmenin en iyi yolu, mümkün olduğunda bireyleri etiketlemekten kaçınmak ve katılımcıların bireyselliklerini koruyacak bir üslup kullanmaktır. Örneğin *bellek kayıplılar* veya *bunaklar* demek yerine "bellek kaybı olan hastalar" veya "bunama grubu" kullanmak daha iyi bir seçenektir. Etiketlenen grup tarafından aşağılayıcı olarak algılanabilecek bir etiket asla kullanılmamalıdır. Bu kılavuzu takip etmeye çalışırken, birey gruplarını etiketlemeyle ilgili tercihlerin zaman içinde değiştiğini ve bir grup içindeki kişilerin, hangi etiketin tercih edileceği konusunda anlaşamayabileceklerini hatırlamak önemlidir.

Örneğin, bazı Kuzey Amerika kökenli bireyler, "Kuzey Amerika Yerlileri" biçiminde adlandırılmayı tercih ediyor olsa da diğerleri "Yerliler" terimini ve başka diğerleri de kendi özgül topluluk adlarını, örneğin Navajo, hatta daha da uygun şekilde, Navajo yerine, kendi ana dillerini kullanarak *Diné*, sözcüğünü istiyor olabilirler.

(c) Bireyler hakkında yazarken, çalışmanızın katılımcılarını açık bir biçimde belirtecek şekilde yazınız. Bunu sağlamanın bir yolu, katılımcıları, *denekler* gibi kişisel olmayan terimler yerine, üniversite öğrencileri ya da çocuklar gibi daha betimleyici terimler kullanarak belirtmektir. Katılımdan bahsederken etken çatı kullanmak, edilgen çatı kullanmaktan daha iyidir. Örneğin, "öğrenciler anketi tamamladı" ifadesi, "anket öğrencilere sunuldu" ifadesinden daha fazla tercih edilmektedir.

- **İLGİ ÇEKİCİ BİR RAPOR YAZIN:** Bilimsel yazının sıkıcı olması gerekmez! Açıktır ki, bilimsel yazı yazanlar bir roman yazarı veya deneme yazarına verilen türden bir lisansa sahip değildirler; ayrıca bilimsel yazı, yaratıcı yazma kursunda öğrendiklerinizi sergileyecek bir yer değildir. Yine de ne yaptığınız, ne bulduğunuz ve nasıl bir sonuca vardığınız konusunda okuyucunuzun ilgilisini çekecek biçimde yazmak için çaba sarf edilmelidir. Araştırma raporlarınızı yazarken uygun bir tarz yakalamaya çalışmanın bir yolu, araştırmanız hakkında iyi bir öykü anlatmaya çalışmanızdır. İyi araştırmalar, iyi öykülerin yolunu açar ve iyi anlatılmış öyküler, araştırmaların ilerlemesini sağlar.

Araştırma raporu yazmak amacıyla hazırlık yaptığınız için, psikoloji alanında ilginizi çeken araştırmaları rapor eden dergi makalelerini okumanızı özellikle öneriyoruz. Bununla birlikte, araştırma raporu yazmak için gerekli olan yetenekleri, ancak sadece raporları gerçekten yazarak geliştireceksiniz.

ARAŞTIRMA RAPORUNUN YAPISI

Bir araştırma raporunun yapısı, yazar ve okuyucu için tamamlayıcı niteliktedir. Söz konusu yapı, yazarın, araştırmaya ait açık betimlemeler ve bulgulara ait ikna edici yorumlar sunmak üzere kullanabileceği bir organizasyon sağlar. Araştırma raporunun okuyucusu, her bir bölümde belirli bilgileri bulmayı bekleyebilir. Bir deneyin nasıl gerçekleştiğini bilmek istiyorsanız Yöntem bölümüne bakarsınız; çalışmadaki verinin analizi ile ilgili bilgi istiyorsanız, Sonuçlar bölümüne başvurursunuz. Bir araştırma raporu aşağıdaki bölümleri içerir:

Başlık Sayfası (Yazar Notu ile birlikte)
Özet
Giriş
Yöntem
Sonuçlar
Tartışma
} **Raporun Ana Gövdesi**
Kaynaklar
Dipnotlar (varsa)
Tablolar ve Şekiller
Ekler (varsa)

Makale Taslağının Biçimine İlişkin İpuçları Tercih edilen yazı karakteri, boşluk verme, aralıklar, paragraf yapılandırması, sayfa numarası verme, başlıkların uygun şekilde kullanımı ve makale taslağı biçimiyle ilgili diğer konular hakkında bilgi edinmek için www.apastyle.org sayfasını ziyaret etmeniz veya *Yayın Kılavuzunu* kullanmanız gerekmektedir.

Başlık Sayfası

Bir araştırma raporunun ilk sayfası başlık sayfasıdır. Bu sayfa araştırmanın ne ile ilgili olduğunu (yani başlık), araştırmayı kimin yaptığını (yani yazarlar), araştırmanın nerede yapıldığını (yani yazarların çalıştığı kurumlar), okuyuculara makalenin ne ile ilgili olduğunu bildirmek için kısa bir başlığı ("kısa başlık") ve yazar notunu gösterir. Yazar notu, yazarın profesyonel mesleğini ve iletişim bilgilerini belirtir; aynı zamanda, varsa teşekkür edilenleri listeler. Bölüm 3'te yazarlığa ait ölçütleri tartışmıştık; sadece kriterlere uyanlar, bir araştırma raporunun yazarları olarak listelenmelidir. Araştırmaya katkıda bulunan diğer kimselere ise yazar notunda teşekkür edilir.

Başlık, okunma olasılığı en yüksek olan bölüm olduğu için, belki de makalenizin en kritik özelliğidir. Anahtar değişkenlerin ve kuramsal meselelerin tanımlanmasıyla başlık, makalenizin merkez konusunu açıkça belirtmelidir. "..... ilişkin bir laboratuvar çalışması" ya da ".....ait bir araştırma" gibi gereksiz sözcüklerden kaçının.

Başlığın Yazılmasına İlişkin İpuçları Bir araştırma raporunun başlığı, araştırma sorusunu veya konu ile ilgili değişkenleri belirtir. Başlık sadece bilgilendirici değil aynı zamanda ilgi çekici de olmalıdır. En önemlisi, başlığınızın, araştırmanızın içeriğini mümkün olduğu kadar özgül bir şekilde betimlediğinden emin olunuz.

Özet

Özet, araştırma raporunun içeriğine ve amacına ait bir paragraflık, kısa bir toparlamadır. Bir özetteki sözcük sayısı sınırlamasıyla ilgili kurallar bilimsel dergilere göre değişmektedir. Bu ilkeler ile ilgili olarak *Yayın Kılavuzu'na* başvurunuz. Görgül bir çalışmanın özeti tipik olarak aşağıdakileri kapsar:

(a) araştırılmakta olan problem;
(b) değişkenlerin nasıl incelendiğini de içeren yöntem;
(c) başlıca bulgular ve
(d) bulgulara ait varılan sonuçlar ve çıkarımlar.

Bir başka deyişle özet, Giriş, Yöntem, Sonuçlar ve Tartışma'da değinilen kritik noktaları vurgulamalıdır. İyi yazılmış bir özet, dergi makalesinin geri kalanının okunup okunmayacağı ile ilgili büyük bir etkiye sahiptir. Özetler, bilgi servisleri tarafından makalelerin indekslenmesi ve edinilmesi amacıyla kullanılmaktadır.

Yayın Kılavuzu, görgül çalışmalar için özetin kritik ögelerini ve literatür gözden geçirmeleri, meta-analizler, kuramsal yazılar, yöntemsel yazılar ve vaka çalışmalarında özetin nasıl farklılaşması gerektiğini daha ayrıntılı olarak betimlemektedir.

Özetin Yazılmasına İlişkin İpuçları İyi bir özet yazımı zorludur. Bu zorlukla baş etmenin en iyi yolu, özeti sonda yazmaktır. Özeti, raporun geri kalanını yazdıktan sonra yazmakla, kendi sözcüklerinizi daha kolay özetleyebilecek veya yorumlayabileceksiniz.

Giriş

Girişe Ait Hedefler Giriş, üç ana hedefe hizmet eder:

1 Üzerinde çalışılmakta olan problemi tanıtmak ve problemin araştırılmasının neden önemli olduğunu belirtmek

2 Çalışma ile ilgili, konuya ait geçmiş literatürü kısaca özetlemek ve çalışmanın kuramsal çıkarımlarını betimlemek ve

3 Araştırmayı yönlendiren tahminler ya da hipotezlerle birlikte, söz konusu çalışmaya ait amaç, gerekçe ve deseni betimlemek.

Makalenizde bu hedeflere değindiğiniz sıra değişebilir, fakat burada açıkladığımız sıra yaygın olarak kullanılanlardan biridir.

Daha önce bahsedildiği gibi giriş, ilgili araştırmaların bir özetini içerir. Bu gözden geçirmenin amacı, literatürün ayrıntılı bir incelemesini vermek değildir. Bunun yerine, araştırmanızla en doğrudan ilişkili çalışmaları dikkatlice seçmelisiniz. Seçilen bu çalışmaları özetlerken, önceki çalışmalara ait, okuyucunun neyi neden yaptığınızı anlamasına yardımcı olacak her ne ayrıntı varsa vurgulamalısınız. Problemi anlamanıza katkı sağlayan diğer araştırmacılardan bahsetmelisiniz. Tabi başka bir kişinin çalışmasından, doğrudan bir alıntı yapıyorsanız tırnak işareti kullanmalısınız (başkalarının çalışmalarına atıf yapma ile ilgili öneriler için bk. Bölüm 3).

Kaynak gösterme genellikle başka araştırmacıların çalışmaları için iki yoldan biri ile yapılır. Ya atıf yaptığınız makalenin yazarlarının soyadlarıyla, soyadlarından hemen sonra parantez içinde makalenin yayımlandığı tarih ile birlikte kaynak gösterirsiniz ya da yazarların çalışmasına genel bir referans verip, sonrasında parantez içinde hem soyadlarını hem de yılları eklersiniz. Örneğin Alice H. Eagly ve Wendy Wood'un 2013 yılında yayımladıkları bir çalışmaya atıf yapacaksanız ya "Eagly ve Wood (2013), etkileşimli bir süreç olarak doğuştan gelen özelliklerin ve eğitimin, cinsiyet farklılıklarının gelişimi üzerindeki etkilerini betimlemişlerdir." ya da "Doğuştan gelen özellikler ve eğitim, cinsiyet farklılıklarının oluşması sırasında etkileşmektedirler (Eagly ve Wood, 2013)." yazarsınız. Eagly ve Wood'un makalesine ait, makale başlığı, dergi adı, cilt numarası ve özgül sayfaları içeren kaynakça bilgilerini, raporun sonunda yer alacak olan Kaynaklar bölümünde tamamlayınız. Dipnotlar, psikoloji alanındaki bir araştırma raporunda kaynaklara atıfta bulunmak için kullanılmamaktadır. Bölüm 3'teki, çalışmanızda kaynaklara atıfta bulunma ile ilgili etik meselelere ait tartışmayı gözden geçirmenizi öneriyoruz (Psikoloji Araştırmalarının Rapor edilmesi başlıklı alt bölüme bakınız).

Özetle, araştırılmakta olan problem, ilgili araştırma bulguları ve çalışmanızın gerekçesi ve deseni, açık, ilgi çekici bir biçimde belirtilmelidir.

Girişin Yazılmasına İlişkin İpuçları Etkili bir giriş yazmak için, yazmaya başlamadan önce neyi neden yaptığınızı kendi kendinize ifade ettiğinizden emin olun. Kendinizi "test" etmenin en iyi yollarından biri, çalışmanıza aşina olmayan birisine çalışmanızın amacını, bu alandaki diğer çalışmalarla olan ilişkisini (örneğin, çalışmanız, zaten bilinen şeylerden hangi yönüyle farklıdır), kuramsal çıkarımlarını ve neye ulaşmayı umduğunuzu sözel olarak betimlemeye çalışmaktır. Dinleyicinizin soruları olabilir, bu soruları yanıtlayarak belki de giriş bölümünüzü yazarken nelerin daha açık hâle getirilmesi gerektiğini bulabilirsiniz.

Psikoloji Literatürünün Taranması Konunuz veya araştırma sorunuz ne olursa olsun, kuşkusuz, psikoloji literatürünü taramaya ihtiyaç duyacağınız bir zaman gelecektir. Örneğin, bir deney ile ilgili fikriniz olsa bile deneyin önceden yapılıp yapılmadığını bulmanız gerekecek. Ya da bir deney için temel almak istediğiniz bir çalışmayı betimleyen bir makale okumuş olabilirsiniz; bu nedenle, giriş yazmak için, ilgili diğer çalışmalar hakkında bilgi edinmek önemli olacaktır. Araştırmanın söz konusu alanı ile ilgili daha fazla şey öğrendikçe, çalışmaya ait önceki fikrinizin gözden geçirilmesi gerektiğini fark edebilirsiniz. Ek okumalar için önemli bir kaynak, konunuz ile ilgili olan makalelerin Kaynaklar bölümüdür.

Psikoloji literatürünü taramak için başlıca çevrimiçi veri tabanı PsycINFO'dur. PsycINFO'ya FirstSearch, InfoTrac gibi çevrimiçi veri tabanlarından ve APA'nın web sayfasından (apa.org) erişilebilir. Yerel kütüphane personeliniz ile birlikte hangi çevrim içi hizmetlerin sizin için ulaşılabilir olduğunu kontrol ediniz. Bir elektronik veri tabanı, veri tabanında yer alan makalelere ait başlıkların ve özetlerin taranmasını ve belirli anahtar sözcükleri içeren makalelerin hepsinin saptanmasını olanaklı kılar. Bu türden bir tarama için en etkili yaklaşım, her ikisinin de, bilgisayar bir makaleyi "işaretlemeden" önce sunulmasının gerekli olduğu, kesişen anahtar sözcükler bulmaktır. Örneğin, bir öğrenci, randevularda gerçekleşen tecavüz ve diğer cinsel saldırı vakalarının (yani ilk randevu tecavüzü – date rape) gerçekleşme sıklığını belirlemek üzere bir tarama tipi araştırma gerçekleştirmek istiyordu. Öğrenci, aramasını yönlendirmek üzere RAPE anahtar sözcüğünü ve devamında DAT metnini kullandı. DAT metnini kullanmasının amacı, DATE, DATES, DATING gibi farklılıkları yakalayabilmekti.

Bu türden çok geniş kapsamlı veri tabanlarının farklı anahtar sözcüklerle defalarca taranmasından sonra, konu hakkında her şeye ulaştığımızla ilgili olarak kendimizden, gereğinden fazla emin olabiliriz. Bununla birlikte, *elektronik bir veri tabanındaki herhangi bir aramada, konu ile ilgili bilginin gözden kaçırılması olasıdır.* Ayrıca anahtar sözcükler aldatıcı olabilir. DAT metni, DATA sözcüğünü kullanan tüm çalışmaları tespit etmiştir, bu nedenle, öğrencinin kaynaklarının bir kısmı tecavüz ile ilgili veri (data) sunan–fakat tamamen randevu (dating) bağlamıyla ilgili olmayan kaynaklardır. Elektronik veri tabanları uygun şekilde kullanıldığı zaman, PsycINFO'yu kullanarak psikoloji literatürünü taramanın avantajları, dezavantajlarına çok daha ağır basacaktır.

Yöntem

Araştırma raporunun ikinci ana bölümü Yöntem bölümüdür. İyi bir Yöntem bölümü yazmak zor olabilir. Yazımı kulağa kolay gelebilir çünkü tek yapmanız gereken, tam olarak araştırmanın nasıl yürütüldüğünü betimlemektir. Fakat bunun ne kadar zor olduğu hakkında bir izlenim edinmek istiyorsanız, sadece ayakkabı bağcığını nasıl bağladığınızı anlatan açık ve ilgi çekici bir paragraf yazmayı deneyin.

Yöntem Bölümünün Yazılmasına İlişkin İpuçları İyi bir yöntem bölümü yazmanın anahtarı organizasyondur. Neyse ki bu bölümün yapısı, farklı farklı araştırma raporları arasında oldukça tutarlıdır, yani birkaç temel alt bölüm, araştırma raporlarının çoğu için ihtiyaç duyduğunuz organizasyonu sağlar. Bununla birlikte, ilk araştırma raporlarını yazan öğrencilerin en sık sorduğu soru şudur: "Ne kadar ayrıntıya inmeliyim?" Çok fazla ya da çok az ayrıntıya girerseniz yazınızın kalitesi olumsuz yönde etkilenecektir. Örneğin, sonuçları kaydetmek için "2 numara kurşun kalem" kullanmış olmanız açıkça gereksiz bir ayrıntıdır! Kabul gören iyi bir ölçüt şudur: Konu ile ilgilenen bir araştırmacının, çalışmanızı tekrarlayabileceği kadar ayrıntıya giriniz. Dergi makalelerinin Yöntem bölümlerini okumak bu yazma görevinde size yardımcı olacaktır. Ayrıca Yöntem bölümünün yazılmasına ilişkin yardım için *Yayın Kılavuzuna*, özellikle ss. 29–32'ye bakınız.

Yöntem bölümünde, çalışmanızda yer alan katılımcıların (deneklerin) sayısını ve niteliğini, kullanılan belirli araçları, aygıtlar ile donanımı ve aynı zamanda çalışmayı tam olarak nasıl yürüttüğünüzü (yani işlem yolunu) betimleyin. Bu türden bilgiler genellikle farklı alt bölümlerde sunulur (örneğin, Katılımcılar, Araçlar, İşlem Yolu) ve bu alt bölümlerin içeriği için APA kılavuzunu gözden geçirmelisiniz. Bu alt bölümlerde neyin yer aldığı konusunda fikir edinmek için, yayımlanmış dergi makalelerinin Yöntem bölümlerini okumak da iyi bir fikirdir.

Sonuçlar

Bu bölüm birçok açıdan, bir araştırma raporunun en heyecanlı kısmıdır çünkü Sonuçlar bölümü araştırma raporunun doruk noktasını içerir –yani çalışmanın asıl bulgularını. Yine de birçok öğrenci için, doruk noktasını betimlemenin heyecanı, istatistiksel bilginin Sonuçlar bölümünde rapor edilmesi gereğiyle ilgili endişe yüzünden körelir. Elbette bu endişeyi azaltmanın en iyi yolu, diğer kavramlarda olduğu gibi, istatistiksel kavramlar için de aynı kuralları uygulamaktır. Yardımcı olabilecek bir ilk adım, Sonuçlar bölümünüzü yazmanıza rehberlik etmesi için basit bir organizasyonel yapı benimsemektir (bk. Tablo 13.1).

Sonuç bölümünü, Giriş bölümünüzde ortaya attığınız soruları yanıtlamak için kullanmalısınız. Bununla birlikte, Sonuçlar bölümündeki temel prensip "olgulara, sadece olgulara odaklanmaktır". Tartışma bölümüne geçtiğinizde, salt olguların ötesine geçme fırsatınız olacak.

İstatistiklerin Rapor Edilmesi Çalışmanızın ham verileri, (örneğin, bireysel puanlar) Sonuçlar bölümünde rapor edilmemelidir. Daha ziyade, özet istatistikleri (örneğin, ortalamalar, standart sapmalar) kullanacak ve hipotezinizle ilgili

TABLO 13.1 SONUÇLAR BÖLÜMÜNE AİT TİPİK BİR PARAGRAFIN YAPISI

1. Analizin amacını belirtin.
2. Sonuçları özetlemek için kullanılacak olan betimsel istatistiği belirtin.
3. Farklı koşullar için bu betimsel istatistiğin bir özetini metin içerisinde, bir tabloda ya da bir şekil ile belirtin.
4. Tablo ya da şekiller kullanıldıysa, okuyucunun odaklanması gereken başlıca bulguları belirtin.
5. Güven aralıklarının, etki büyüklüklerinin ve çıkarımsal istatistiklerin kullanılma nedenlerini ve sonuçlarını belirtin.
6. Her testten sonra, varılan sonucu belirtin fakat çıkarımları tartışmayın. Bunlar Tartışma bölümüne aittir.

Örnek paragraf

Çalışma sırasında verilen yönergelerin bir fonksiyonu olarak bellekte tutma miktarını incelemek amacıyla, her katılımcı tarafından, her yönerge koşulunda anımsanan sözcük sayısı belirlenmiştir. Sözcükler, sadece hedef listede görünmüş olan bir sözcükle eşleştiği durumda "doğru" olarak puanlanmıştır. Anımsanan ortalama sözcük sayısı (standart sapmalarla birlikte): tuhaf imgeleme koşulu, standart imgeleme koşulu ve kontrol koşullarında sırasıyla 15.6 (1.44), 15.2 (1.15) ve 10.1 (1.00)'dir. %95'lik GA'ları: Tuhaf imgeleme [13.18, 18.02], standart imgeleme [12.78, 17.62], kontrol [7.68, 12.52]. Genel olarak, ortalamalar arasındaki farklar istatistiksel açıdan anlamlıdır $F(2,72) = 162.84$, $p < .001$, HOK = 1.47, $\eta^2 = .82$. Güven aralıklarının karşılaştırılması, her iki imgeleme koşulunun da kontrol koşulundan farklılaştığını göstermiştir fakat iki imgeleme koşulu farklılaşmamaktadır. Sonuç olarak, imgelemeyi kullanmaları yönergesi verilen katılımcıların bellekte tutma miktarları, belirli bir çalışma yönergesi verilmeyen katılımcılara göre daha yüksektir fakat bellekte tutma, iki tür imgeleme yönergesi arasında farklılaşmamaktadır.

olan çıkarımsal testlerin sonuçlarını (hem beklediğiniz, hem de beklemediğiniz sonuçları!) rapor edeceksiniz. Karmaşık çalışmalar için, tablo ve şekillerin kullanımı genellikle önemlidir. Sonuçlar bölümü, Tartışma bölümünde rapor edeceğiniz sonuçlar için zemin hazırlar.

İyi Bir Sonuçlar Bölümünün Yazılmasına İlişkin İpuçları Sonuçlar bölümünüzü yazarken şu adımları takip etmenizi öneriyoruz.

- *Adım 1*: Bir sonuç bölümü paragrafı, analizin amacının belirtilmesiyle başlar. Analizi yapmanın neden(leri)i kısa ve öz bir biçimde ifade edilmelidir; genellikle bir cümleden fazlası gerekli değildir. Örneğin, analizin amacı "çalışma sırasında verilen yönergelerin bir fonksiyonu olarak bellekte tutma miktarını incelemek"tir.
- *Adım 2*: Sonuç bölümü paragrafının yazımındaki ikinci adım, belirli bir bağımlı değişkene ait sonuçları özetlemek için kullanılacak olan betimsel istatistikleri belirtmektir (örneğin; ortalama, ortanca, toplam frekans). Örnek paragrafta araştırmacılar sonuçları özetlerken, anımsanan sözcük sayısı ortalamasını kullanmışlardır.

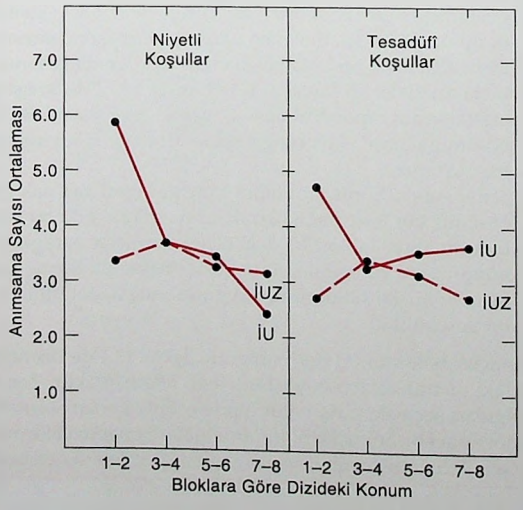
- *Adım 3*: Üçüncü adım, bu betimsel istatistiklerin koşullara göre bir özetinin sunulmasıdır. Ortalama ile birlikte standart sapmanın rapor edilmesinde olduğu gibi, merkezî eğilim ölçülerine, konu ile ilgili değişkenlik ölçüleri eşlik etmelidir. Bir etki büyüklüğü ölçümünün kullanılması da özellikle önerilmektedir. Deneyinizde sadece iki veya üç koşul söz konusu ise, bu özet, metin içinde sunulabilir. Fakat özetleyecek daha fazla veriniz varsa bulgularınızı ya bir tabloda ya da bir şekilde (grafik) sunmanız gerekir. Tabloların ve şekillerin yapılandırılması ile ilgili işlemleri bu bölümde daha sonra betimleyeceğiz.
- *Adım 4:* Bir tablonun ya da şeklin tek başına yeterli olduğu düşünülmemelidir. Okuyucunuzun, bir tablo veya şekilden mümkün olduğu kadar çok bilgi elde etmesi için yardıma ihtiyacı olacaktır. Sonuçlarınıza en aşina kişi siz olduğunuz için, bu yardımı verebilecek en uygun pozisyonda bulunuyorsunuz. Tabloda veya grafikte, verinin ilgi çekici kısımlarına, özellikle de verinin, girişte önerdiğiniz hipotezler ile uyumlu (veya uyumsuz) taraflarına okuyucularınızın dikkatini yönlendirmelisiniz. Genellikle aynı veri hem tabloda hem de şekilde sunulmamaktadır. Hangisini seçerseniz seçin, tablo ya da şeklin ortaya koyduğu kritik sonuçları metin içinde vurguladığınızdan emin olunuz.
- *Adım 5:* Sonuçlar bölümüne ilişkin bir paragrafın yazılmasındaki beşinci adım, çıkarımsal istatistik testlerine ait sonuçların sunulmasıdır. Herhangi bir istatistiksel test için, şu bilgiler her zaman rapor edilmelidir: testin adı (genellikle *t*, *r*, *F* gibi bir sembol ile gösterilir); teste ait serbestlik derecesi (test yazıldıktan sonra parantez içinde sunulur); elde ettiğiniz test istatistiğinin değeri; test sonucuna ait tam olasılık (örnek paragrafta olduğu gibi *p* değeri .001'den küçük olmadıkça) ve etki büyüklüğü ölçüleri. Ayrıca Tablo 13.1'de (örnek paragrafa bakınız) gösterildiği gibi hata ortalama kareyi de (*HOK*) eklemelisiniz. *HOK* (*F* oranının paydası), ilgilenen okuyucuların, sonuçlarınızdan hareketle ek istatistikleri hesaplamasına izin verir ve daha sonra yapılacak olan meta-analizleri kolaylaştırır. Bölüm 11 ve 12'de tartıştığımız gibi, güven aralıklarının rapor edilmesini önemle öneriyoruz. Bu bilginin, Sonuçlar bölümüne nasıl ekleneceğine dair örnekler için yine örnek paragrafa başvurunuz.
- *Sonuç Adımı:* Sonuçlar bölümünde paragraf yazmak için son adım, rapor ettiğiniz her teste bağlı olarak ortaya çıkan, kısa bir sonuç ifadesinde bulunmaktır. Tablo 13.1'deki örnek Sonuçlar paragrafındaki son cümle, bunun nasıl yapılabileceğini göstermektedir. Sonuç apaçık ortada olabilir fakat Sonuçlar bölümünde uygun sonuç ifadelerinin kullanılması temel bir gerekliliktir.

Sonuçlar bölümünün her paragrafı, Tablo 13.1'de özetlenen yapıyı izler. Buradaki mantık, okuyucuyu istatistiğe boğmamaktır. Zor olan şey, en önemli bulguları seçmek, giriş bölümünde ortaya konan sorulara ilişkin tüm veriyi rapor etmektir. Sonuçlar bölümüne ait tartışmamızı bitirmeden önce, tabloların ve şekillerin oluşturulmasına ait temel işlemleri kısaca betimleyeceğiz.

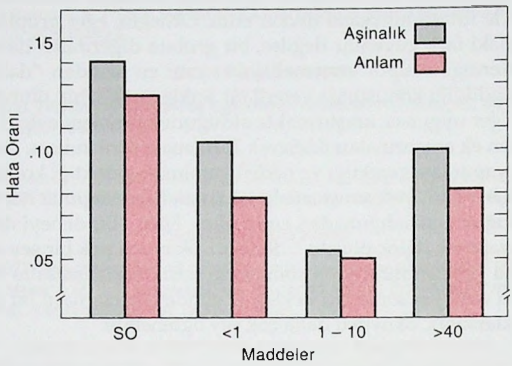
Verinin Tablolarla Sunulması Tablolar, büyük miktarlardaki veriyi kısa bir biçimde sunmanın etkili ve verimli bir yoludur. Tablo, makale metnindeki bilgiye destek olmalı ve bu bilgiyi tekrarlamamalıdır fakat metinle de iyi bir şekilde bütünleşmelidir. Bir araştırma raporundaki tablolar, ardışık olarak numaralandırılır. Tabloları numaralandırmak, metinde numaraları kullanarak tablolara gönderme yapmayı kolaylaştırır. Ayrıca her tablonun kısa bir açıklayıcı başlığı olmalıdır ve sütunlar ile satırlar açık bir şekilde isimlendirilmelidir. Tablodaki veri girdilerinin hepsi aynı hassasiyetle rapor edilmelidir (yani, tüm değerler aynı sayıda ondalık hanesine sahip olmalıdır) ve değerler ilgili satır ve sütun başlıklarına uyumlu bir şekilde hizalanmalıdır. APA stili gereklerine göre tabloların farklı hazırlanma şekillerini görmek için *Yayın Kılavuzuna* başvurmanız gerekir (*Kılavuzda* özellikle Bölüm 5'e bakınız).

Verinin Şekillerle Sunulması Şekiller, tablolar gibi, büyük miktarlardaki verileri sunmanın kısa bir yoludur. Bir şeklin iki temel ekseni vardır: yatay eksen, yani x-ekseni ve dikey eksen, yani y-ekseni. Genellikle, bağımsız değişkenin düzeyleri x-ekseninde ve bağımlı değişkeninkiler ise y-ekseninde gösterilir. İki ve daha fazla bağımsız değişken varsa, ikinci ve daha sonra gelen bağımsız değişkenler, şeklin içindeki etiketler şeklinde yer alır ya da şekil açıklamasında gösterilir. Şekil 13.1'de, bağımlı değişken değerleri (ortalama anımsama sayısı) y-ekseninde, bağımsız değişkenin düzeyleri (dizisel konum) x-ekseninde gösterilmiştir. İkinci bağımsız değişkenin düzeyleri [ipuçlu (İU) ya da ipuçsuz

ŞEKİL 13.1 Bloklara, ipucuna (İU = İpuçlu; İUZ = İpuçsuz) ve yönerge koşuluna göre, dizideki konumun bir fonksiyonu olarak, anımsanan ortalama sözcük sayısı (mümkün olan 10 tanesinin arasından)



ŞEKİL 13.2 Sözel maddelerin aşinalık veya anlam açısından puanlanmasından sonra, iki grup üniversite öğrencisi tarafından yapılan tanıma hatalarının oranı. Maddeler, sözcük olmayanlar (SO) ve Thorndike-Lorge sayımında her bir milyonda, 1 defadan az, 1 ile 10 defa arasında ve 40 defadan fazla görünen sözcüklerdir.



(İUZ)] figür içindeki veriyi adlandırmaktadır ve üçüncü bağımsız değişkenin düzeyleri (yönergeler) şeklin iki ayrı bölümü için başlık işlevi görmektedir.

Psikolojide iki genel şekil türü yaygın olarak kullanılmaktadır: Çizgi grafikler ve çubuk grafikler. En yaygın şekil türü, Şekil 13.1'de görüldüğü gibi, çizgi grafiktir. Bununla birlikte *x*-ekseninde gösterilen bağımsız değişken sınıflayıcı ölçek düzeyindeki bir değişken ise genellikle çubuk grafik kullanılmaktadır. Örneğin farklı bölümlere (bağımsız değişken) kayıtlı öğrencilerin GNO ortalamasını (bağımlı değişken) gösteriyorsanız çubuk grafik kullanabilirsiniz. Çubuk grafiğe ait bir gösterim Şekil 13.1'dedir. Kullanışlı grafiksel sunumlar oluşturmak için alternatif yollar vardır ve çeşitli seçenekleri görmek için *Yayın Kılavuzuna* (Bölüm 5) başvurun. Tüm şekiller, anlaşılır bir biçimde çizilmeli ve uygun biçimde adlandırılmalıdır; böylece okuyucular tam olarak neyin sunulduğunu anlayabileceklerdir.

Tartışma

Tartışma bölümü, Sonuçlar bölümünden farklı olarak, "olgulardan daha fazlasını" içerir. Bu bölüm, araştırmanızın çıkarımlarını açıklamanın, hipotezinizi destekleyen belirli sonuçları vurgulamanın ve desteklemeyen herhangi bir sonuç üzerinde eleştirel yorum yapmanın yeridir. Başka bir deyişle, okuyuculardan oluşan jüriye son bir özetleme yaparsınız.

Tartışma, temel bulguların kısa ve öz bir ifadesiyle başlar. Bu özette betimsel istatistikleri tekrarlamazsınız; bulgulara ait istatistiksel analizlere değinmenizin de gereği yoktur. Sonuçları, bu alandaki başka kişilerin, özellikle daha önce giriş bölümünde atıfta bulunduklarınızın, bulguları ile karşılaştırmanız gerekmektedir. Okuyucunuza karşı "dürüst" olun ve deseninizdeki ya da analizinizdeki, farklı yorumlara neden olabilecek, eksik yönleri açıkça belirtin. Sınırlılıkları veya sorunları belirtmenin iyi bir yolu, diğerlerinin çalışmanız

hakkında yapabileceği eleştirileri tahmin etmeye çalışmaktır. Eğer sonuçlarınız orijinal hipotezinizle tutarlı değilse, bu uyumsuzluğa ilişkin bir açıklama önermelisiniz.

Tartışmada kullandığınız ifadelerin, Sonuçlar bölümünde rapor edilen verilerle tutarlı olmasına dikkat edin. Örneğin, eğer grupların ortalamaları arasındaki fark güvenilir değilse, bir grubun diğerinden daha iyi bir performans gösterdiğini rapor etmemelisiniz –yani en azından "daha iyi" derken neyin kastedildiği konusunda yeterli bir açıklamaya sahip olunmadığı sürece.

Eğer uygunsa, araştırmakta olduğunuz problemle ilgili olarak yapılması gereken ek araştırmaları önererek Tartışmayı sonlandırın. Ne türden araştırmaların yapılması gerektiği ve neden yapılması gerektiği konusunda özgül olmaya çalışın. Yani, yeni araştırmaların, şu an bilmediğimiz neyi ortaya koyması gerektiğini açıkladığınızdan emin olun. Eğer, "Bu deneyi daha genç katılımcılar ile yapmak ilginç olurdu." derseniz okuyucu pek bir şey anlamayacaktır. Eğer daha genç katılımcılarla sonuçların nasıl farklılaşmasını beklediğinizi ve önerilen deneyin sonuçları beklenen yönde çıkarsa nasıl bir çıkarım yapacağınızı açıklarsanız, okuyucu daha çok şey öğrenebilir.

Tartışma Bölümünün Yazılmasına İlişkin İpuçları Tartışma bölümü için olası bir özet aşağıdaki gibidir:

- Problemin ve hipotezlerin (beklentilerin) kısaca gözde geçirilmesi;
- Hipotezinizi destekleyen (veya desteklemeyen) ana bulguların bir özeti;
- Çalışmanızın sınırlılıkları ile ilgili yorumlar (her zaman bazı sınırlılıklar vardır!);
- Gelecekteki çalışmalar için öneriler (özgül olunuz!);
- Bulguların önemine ilişkin yorumlar ve uygunsa olası uygulamalı etkiler.

Kaynaklar

Araştırma raporlarının çoğunda genellikle dört tür kaynak bulunur: Dergi makaleleri, kitaplar, editörlü kitaplarda bölümler ve internet kaynakları. Tablo 13.2, bir makale taslağının Kaynaklar bölümünde bu kaynaklara nasıl atıf yapılacağını göstermektedir. Bu kaynakların ve daha birçok kaynak türünün rapor edilmesi sırasında kullanılacak, APA stiline göre belirli biçimlendirme kuralları en iyi şekilde, *Yayın Kılavuzuna* başvurularak incelenebilir. www.apastyle.org adresinde yer alan ücretsiz eğitim de ayrıca kaynakları biçimlendirirken size yardımcı olabilir.

Elektronik yayıncılığın hızla yaygınlaşması, internetten elde edilen bilgilere ait elektronik "tanımlayıcılara" gereksinim duyulmasına neden olmuştur. Örneğin, İnterneti kullanan herhangi biri, URL'lere (uniform resource locator) aşinadır. Bunlar genellikle "http://" ile başlamakta ve bir ana makine adı, (genellikle www ile başlar), dosya yolu ve belgenin başlığı ile devam etmektedir. Örneğin, psikoloji alanındaki konularla ilgili araştırmaları bulmanıza ("Psikolojide Kütüphane Araştırması") son derece yardımcı olacak bir çevrim içi kaynak şudur: http://www.apa.org/education/undergrad/library-research.aspx

TABLO 13.2 KAYNAKLARIN GÖSTERİLME ŞEKİLLERİ

DOI'siz Dergi Makalesi

Loftus, E. F., ve Burns, T. E. (1982). Mental shock can produce retrograde amnesia. *Memory & Cognition, 10,* (4), 318–323.

DOI'li Dergi Makalesi

Hyde, J. S. (2005). The gender similarities hypothesis. *American Psychologist, 60,* 581–592. doi:10.1037/0003-066X.60.6.581

Kitap

Posavac, E. J. (2011). *Program evaluation* (8. baskı). Englewood Cliffs, NJ: Prentice Hall.

Editörlü Kitapta Bölüm

Weiss, J. M. (1977). Psychological and behavioral influences on gastrointestinal lesions in animal models. D. Maser & M. E. P. Seligman (Ed.), *Psychopathology: Experimental models içinde* (pp. 232–269). San Francisco, CA: W. H. Freeman.

Çevrimiçi Elde Edilen Teknik Rapor veya Araştırma Raporu

Lenhart, A., Madden, M., ve Hitlin, P. (2005). *Teens and technology: Youth are leading the transition to a fully wired and mobile nation.* Temmuz 2005 tarihinde http://www.pewinternet.org/pdfs/PIP_Teens_Tech_adresinden indirildi.

İnternetten elde edilen bilgiye atıf yapmanız gerekirse, kaynağı konumlandırmak için gerekli olan özgül bilgiyi vermeniz önemlidir.

Elektronik kimlikleyicinin daha yakın zamanlardaki bir biçimi dijital nesne tanımlayıcısıdır (DOI). DOI, bir makalenin ya da internetteki başka bir bilgi kaynağının içeriğini ve elektronik konumunu belirten alfanumerik bir metindir. DOI (ya da küçük harflerle doi) genellikle yayımlanmış bir makalenin başlık sayfasında bulunur. APA stili kuralları, DOI mevcutsa, atfınızın bir parçası olarak Kaynaklar bölümünde bunu eklemenizi gerektirir. DOI'yi kullanmanın kolay bir yolu, arama yaparken DOI'yi, http://dx.doi.org/ adresinden sonra eklemektir. Böylece Tablo 13.2'deki 10.1037/0003-066X.60.6.581 tanımlayıcısına sahip makale, arama motorunuzda http://dx.doi.org/10.1037/0003-066X.60.6.581 adresini kullanarak bulunabilir. Bu DOI'yi kullanmayı ve Hyde (2005) kaynağını bulup bulamadığınızı görmeyi deneyin. Bir defa daha sizi, elektronik kaynaklara ve önerilen kaynak biçimlerine ait daha geniş kapsamlı bir tartışma için *Yayın Kılavuzuna* yönlendiriyoruz.

Kaynak biçimlerini yakından takip ederseniz ve kaynak listenizi dikkatlice gözden geçirirseniz, okuyucularınızın işini kolaylaştırmış olursunuz. Kaynaklar, her makalenin ilk yazarının soyadına göre alfabetik sıra ile listelenmektedir.

Dipnotlar

Dipnotlar, dergi makalelerinde nadiren bulunur, öğrencilerin araştırma raporlarında daha da nadirdir. Kullanıldıkları zaman, metinde ardışık olarak numaralandırılmalıdır ve Kaynaklar bölümünden sonra ayrı bir sayfaya konmalıdır.

Ekler

Yayımlanmış araştırma makalelerinde Ekler nadiren bulunur ama öğrencilerin araştırma raporlarında daha yaygındır. Yayımlanmış bir makale için kullanılacaksa her ek, makale taslağının ayrı bir sayfasından başlar ve yazının sonunda kaynaklardan sonra görünür (Not: Öğretim elemanları, ham verinizi, istatistiksel bir analize ait çalışma sayfalarını ya da analizlerin bilgisayar çıktısını içeren bir ek göndermenizi isteyebilir. Ek ayrıca, katılımcılara verilen yönergelerin tam bir kopyasını veya bir deneyde kullanılan belirli araçların listesini sunmak için kullanılabilir.) Her ek, bir harf (A, B, C vb.) ile tanımlanır ve metnin gövdesinde eke yapılacak herhangi bir gönderme bu harf kullanılarak yapılır. Örneğin "Yönergenin tamamı Ek A'dadır" şeklinde yazabilirsiniz.

Makale Taslağınızı Dergi Editörüne Göndermeye İlişkin İpuçları *Yayın Kılavuzu*, editöryal ilkelerin betimlemesini, yazar sorumluluklarını, makale taslağı kontrol listesini, dergi editörü için örnek açıklayıcı mektubu ve APA dergilerine makale taslaklarını gönderirken gerekli olabilecek Etik İlkeler Formu ile birlikte APA Uygunluğunu da içeren yayın süreci ile ilgili önemli bilgileri sunmaktadır (*Kılavuzun* Ekindeki Tablo 1, 2 ve 3, makale taslaklarına dâhil edilmesi önerilen, orijinal veri toplama sürecinin rapor edilmesi ile ilgili geniş kapsamlı bilgi içermektedir.

SÖZEL SUNUMLAR

Araştırma psikologları düzenli olarak, araştırmalarının kısa bir sözel betimlemesini sundukları profesyonel konferanslara ve toplantılara katılmaktadırlar. Benzer şekilde, öğrenciler de sınıfta ya da birkaç farklı sınıftan öğrencinin katıldığı bir bölüm araştırma sempozyumunda veya lisans düzeyindeki araştırma konferanslarında araştırmalarının sözel sunumunu yapabilirler. Tüm bu ortamlar ortak bir özelliği paylaşmaktadır –sunum için ayrılan zaman genellikle 10, 15 dakikadan fazla değildir. Bu zaman aralığında, dergi makalesindeki gibi ayrıntılı bir betimleme yapmak olanaksızdır.

İyi bir sözel sunum, probleme, yönteme, ana bulgulara ve varılan sonuçlara ait az ve öz bir genel bakış sağlar. Birçok açıdan, çalışmanızın genişletilmiş bir özetine benzemektedir. Araştırmacılar sıklıkla, sözel sunumda verilebileceğinden daha fazla ayrıntı içeren, çalışmalarına ait yazılı bir kopya vermektedir. Bu durum, sunum yapan kişinin, çalışmanın önemli noktalarına değinmesine ama yöntem veya analizlerin ince ayrıntılarına girmemesine olanak sağlar. Özgül istatistiksel sonuçları rapor etmekten kaçının ("ANOVA sonucunda elde edilen *F* değeri, 4.67'dir"). Basitçe, "anlamlı bir fark bulundu" ya da "koşullar güvenilir şekilde farklılaştı" şeklinde rapor edin. Dinleyiciler ayrıntılara yazılı el çıktılarınızdan bakabilirler.

Etkili Bir Sözel Sunumun Yapılmasına İlişkin İpuçları Söz konusu olan şeyin "sözel" sunum olması, bunun tam bir yazılı versiyonunu hazırlamayacağınız anlamına gelmez. Bununla birlikte, konuştuğunuz gibi yazmaya ve örneğin, dergi makalesi yazıyor gibi yazmamaya dikkat ediniz. Basit cümleler kullanın ve duraklamak ya da bir görsele başvurmak istediğiniz yerleri işaretleyin. Birçoğumuz gergin olduğumuz zaman daha hızlı konuşuruz, bu nedenle sayfalarınızdaki duraklama işaretleri size makul bir hızda konuşmanız gerektiğini ve ara sıra duraklamanız gerektiğini hatırlatacaktır. Konuşma için kullanacağınız yazılı versiyonun, dinleyicilerinize dağıtılan el çıktıları ile aynı olması gerekmez (hatta aynı olmamalıdır). Sunumunuzu yapmadan önce, belki PowerPoint görsellerindeki yardımcı ipuçlarıyla sunumu hatırlamak veya sunumu okumak size kalmış bir şeydir. Topluluk önünde konuşmak konusunda tedirginseniz, sunumu okumaktan çekinmeyiniz. Çok sınırlı zaman aralıkları söz konusu olduğu için, sunumunuz sadece çalışmanızın önemli noktalarına değinmelidir. Yazılı sunumunuzun hazır olduğundan emin olduğunuzda, bir sonraki adım kendi kendinize yüksek sesle sunumun provasını yapmaktır; böylece, ne söyleyeceğiniz ve zaman sınırınıza uyup uyamayacağınız konusunda fikir edinirsiniz. Sonrasında, eleştirel (fakat dostça) bir topluluğun önünde konuşmanızın alıştırmasını yapmanız gerekebilir. Alıştırma yaptığınız dinleyicilere neyi anlamadıklarını veya neyin daha anlaşılır yapılması gerektiğini sorun. Ne yaptığınızı ve ne bulduğunuzu takip edebildiler mi? Yeteri kadar yüksek sesle konuştunuz mu? Görselleriniz (eğer varsa) açık ve etkili mi? Vurguladığınız başlıca noktaları tekrar edebiliyorlar mı? İzin verilen zaman sınırınıza uyabildiniz mi? Son olarak, "gerçek" bir dinleyici kitlesi önünde sunumunuzu yaparken, sorular için zaman ayırdığınızdan emin olun.

ARAŞTIRMA ÖNERİLERİ

Bu bölümün son kısmında, yine yazım ile ilgili, fakat bu kez araştırma önerilerini, tartışacağız. Bu bölümün başlangıcında sözünü ettiğimiz gibi, araştırmacıların genellikle araştırmaları için, özel kuruluşlara veya devlet kuruluşlarına ödenek önerileri sunarak, finansal destek aramaları gerekmektedir. Öğrencilerin de bazen araştırma yöntemleri dersinde yapabilecekleri araştırmalarını betimleyen öneriler sunması gerekmektedir. Yazılı bir öneri zorunlu olmasa bile, ilgili literatürün, olası pratik problemlerin, veriye ait uygulanabilir istatistiksel analizlerin ve beklenen sonuçların nihaî yorumunun dikkatlice yapılmış bir ön değerlendirmesi olmadan, ancak gözü kara bir araştırmacı bir araştırma projesiyle uğraşır. Bu dikkatli ön değerlendirme size, gerçekleştirilebilir, uygun biçimde analiz edilebilir ve yorumlanabilir bir araştırma projesi geliştirmede yardımcı olacaktır.

Bir araştırma projesinin amacı, kayda değer bir bilimsel yeterliğe dayalı, yorumlanabilir görgül bulgular ile sonuçlanacak, işe yarar bir araştırma desenini garanti etmektir. Hiçbir araştırma önerisi, ne kadar dikkatli hazırlandığından bağımsız olarak, önemli sonuçların elde edileceğini garanti etmez. Araştırmacılar Murphy Yasası'nı kariyerlerinin erken aşamalarında öğrenmektedirler. Esas itibariyle Murphy Yasası şöyledir: "Yanlış gidebilecek herhangi bir şey, yanlış gidecektir." Yine de, önlenebilecek olan araştırma sorunları önlenebilirse, araştırma önerisi geliştirmek yararlıdır.

Yazılı bir araştırma önerisi, genel bir dergi makalesinin şeklindedir fakat çeşitli bölümlerin başlıkları biraz farklıdır. Öneri, şu ana bölümleri içermelidir:

Giriş
Yöntem
Beklenen Sonuçlar ve Önerilen Veri Analizi Planı
Varılan Sonuçlar
Kaynaklar
Ekler
Etik Kurul İçin Bilgi

Araştırma önerisine özet eklenmemektedir. Araştırma önerisinin girişinin, dergi makalesi için gerekli olandan daha geniş kapsamlı bir literatürü incelemesi mümkündür. Bir araştırma önerisinde, araştırma probleminin belirtilmesi ve hipotezlerin mantıksal inşası, dergi makalesinde gerekli olanla aynıdır. Benzer biçimde, önerideki Yöntem bölümü, tamamlanmış araştırmadakine mümkün olduğu kadar yakın olmalıdır.

Önerinin "Beklenen Sonuçlar ve Önerilen Veri Analizi Planı" başlıklı bölümü, araştırmanın beklenen sonuçlarının kısa bir tartışmasını içermelidir. Çoğu durumda sonuçların tam olarak nasıl çıkacağı bilinemeyecektir. Yine de, araştırmanın sonucuna ilişkin her zaman bir fikriniz (bir hipotez ya da tahmin biçiminde) olacaktır. Beklenen Sonuçlar bölümü, sonuçlara ait beklediğiniz (umduğunuz) yönde çıkacak tabloları veya şekilleri içerebilir. Beklenen sonuçların proje için en önemli olanları vurgulanmalıdır. Beklenen sonuçlar için önerilen veri analizi planı bu bölümde olmalıdır. Örneğin, karmaşık bir desen öneriyorsanız, hangi etkileri test edeceğinizi ve hangi istatistiksel testleri kullanacağınızı belirtmeniz gerekir. Beklenen sonuçların akla yatkın alternatiflerine ve ayrıca, sonuçlar araştırma hipotezinden saparsa ortaya çıkacak olası yorumlama sorunlarına da değinilmelidir. Araştırma önerisinin gövdesi, beklenen sonuçlar temelinde varılan sonuçların ve çıkarımların kısa bir açıklamasını sunan, Varılan Sonuçlar bölümü ile sonlanmaktadır.

Kaynaklar bölümü, raporun göndereceğiniz son hâliyle tam olarak aynı biçimde olmalıdır. Ekler, araştırma önerisini tamamlamalı ve çalışmayı gerçekleştirirken kullanılacak olan tüm araçların listesini içermelidir. Örneğin, öğrencilerin, sözcük listeleri ile ilgili belleğini inceleyen bir deney yapıyorsanız, asıl sözcük listeleri ve yönergeler eklenmelidir.

Son olarak bir araştırma raporu, önerilen araştırmayı etik açıdan değerlendirmek için kurulmuş bir etik kurula veya benzer bir heyete gönderilmek üzere bir belge içermelidir (bk. Bölüm 3). Hiç şüphesiz kurumunuzun, önerinizle birlikte gönderilmek üzere standart formları vardır.

Ekler

İstatistik Tabloları

EKLERİN ANA HATLARI

TABLO A.1 SEÇKİSİZ SAYILAR TABLOSU
TABLO A.2 *t* DAĞILIMI DEĞERLERİ
TABLO A.3 *F* DAĞILIMINA AİT KRİTİK DEĞERLER

TABLO A.1 SEÇKİSİZ SAYILAR TABLOSU*

Sütun / Satır	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
1	10480	15011	01536	02011	81647	91646	69179	14194	62590	36207	20969	99570	91291	90700
2	22368	46573	25595	85393	30995	89198	27982	53402	93965	34095	52666	19174	39615	99505
3	24130	48360	22527	97265	76393	64809	15179	24830	49340	32081	30680	19655	63348	58629
4	42167	93093	06243	61680	07856	16376	39440	53537	71341	57004	00849	74917	97758	16379
5	37570	39975	81837	16656	06121	91782	60468	81305	49684	60672	14110	06927	01263	54613
6	77921	06907	11008	42751	27756	53498	18602	70659	90655	15053	21916	81825	44394	42880
7	99562	72905	56420	69994	98872	31016	71194	18738	44013	48840	63213	21069	10634	12952
8	96301	91977	65463	07972	18876	20922	94595	56869	69014	60045	18425	84903	42508	32307
9	89579	14342	63661	10281	17453	18103	57740	84378	25331	12566	58678	44947	05585	56941
10	85475	36857	53342	53988	53060	59533	38867	62300	08158	17983	16439	11458	18593	64952
11	28918	69578	88231	33276	70997	79936	56865	05859	90106	31595	01547	85590	91610	78188
12	63553	40961	48235	03427	49626	69445	18663	72695	52180	20847	12234	90511	33703	90322
13	09429	93969	52636	92737	88974	33488	36320	17617	30015	08272	84115	27156	30613	74952
14	10365	61129	87529	85689	48237	52267	67689	93394	01511	26358	85104	20285	29975	89868
15	07119	97336	71048	08178	77233	13916	47564	81506	97735	85977	29372	74461	28551	90707
16	51085	12765	51821	51259	77452	16308	60756	92144	49442	53900	70960	63990	75601	40719
17	02368	21382	52404	60268	89368	19885	55322	44819	01188	65255	64835	44919	05944	55157
18	01011	54092	33362	94904	31273	04146	18594	29852	71585	85030	51132	01915	92747	64951
19	52162	53916	46369	58586	23216	14513	83149	98736	23495	64350	94738	17752	35156	35749
20	07056	97628	33787	09998	42698	06691	76988	13602	51851	46104	88916	19509	25625	58104
21	48663	91245	85828	14346	09172	30168	90229	04734	59193	22178	30421	61666	99904	32812
22	54164	58492	22421	74103	47070	25306	76468	26384	58151	06646	21524	15227	96909	44592
23	32639	32363	05597	24200	13363	38005	94342	28728	35806	06912	17012	64161	18296	22851
24	29334	27001	87637	87308	58731	00256	45834	15298	46557	41135	10367	07684	36188	18510
25	02488	33062	28834	07351	19731	92420	60952	61280	50001	67658	32586	86679	50720	94953
26	81525	72295	04839	96423	24878	82651	66566	14778	76797	14780	13300	87074	79666	95725
27	29676	20591	68086	26432	46901	20849	89768	81536	86645	12659	92259	57102	80428	25280
28	00742	57392	39064	66432	84673	40027	32832	61362	98947	96067	64760	64584	96096	98253
29	05366	04213	25669	26422	44407	44048	37937	63904	45766	66134	75470	66520	34693	90449
30	91921	26418	64117	94305	26766	25940	39972	22209	71500	64568	91402	42416	07844	69618

Basit Bir Seçkisiz Örneklem Elde Etmek İçin Seçkisiz Sayılar Tablosunu Kullanın

1. Örneklem havuzundaki her bir elementi numaralayın.
2. Örneklemin büyüklüğüne karar verin
3. Tablodan, tesadüfi bir şekilde bir başlangıç noktası seçin (örneğin, gözlerinizi kapatarak kaleminizin ucuyla tabloda bir yere dokunun).
4. Yanlamasına (veya aşağı veyahut yukarı) giderek sayıları belirleyin. Not: Bir sayıdaki rakamların sayısı, örneklemdeki rakamların sayısına karşılıktır (örneğin, bir örneklemin büyüklüğünün 147 olması, sizin 001 ile 147 arasında 3 rakamlı sayılar aradığınız, böyle olmayanları ise atladığınız anlamına gelir.
5. İstediğiniz örnekleme ulaşıncaya kadar devam edin. Daha sonra, seçtiğiniz rakamların karşılığı olan elementleri listeleyin.

* Kaynak: Table of 105,000 Random Decimal Digits, Statement no. 4914, File no. 261-A-1, Interstate Commerce Commission, Washington, D.C. May 1949.

TABLO A.2 TABLO A.2 *t* DAĞILIMINDAN SEÇİLMİŞ DEĞERLER*

Kullanma yönergesi: Bir *t* değeri bulmak için, tablonun sol sütununda, ortalamanın standart hatasıyla bağlantılı olan serbestlik derecesine (sd) karşılık gelen sayıyı bulun. Daha sonra, seçtiğiniz çift yönlü *t* değerine uygun *t* değerini seçin. $\alpha = .05$ sütununda verilen değer, %95 güven aralığının hesaplanmasında kullanılır; $\alpha = .01$ sütununda verilen değer ise %99 güven aralığının hesaplanmasında kullanılır.

sd	$\alpha = .05$	$\alpha = .01$	*sd*	$\alpha = .05$	$\alpha = .01$
1	12.71	63.66	18	2.10	2.88
2	4.30	9.92	19	2.09	2.86
3	3.18	5.84	20	2.09	2.84
4	2.78	4.60	21	2.08	2.83
5	2.57	4.03	22	2.07	2.82
6	2.45	3.71	23	2.07	2.81
7	2.36	3.50	24	2.06	2.80
8	2.31	3.36	25	2.06	2.79
9	2.26	3.25	26	2.06	2.78
10	2.23	3.17	27	2.05	2.77
11	2.20	3.11	28	2.05	2.76
12	2.18	3.06	29	2.04	2.76
13	2.16	3.01	30	2.04	2.75
14	2.14	2.98	40	2.02	2.70
15	2.13	2.95	60	2.00	2.66
16	2.12	2.92	120	1.98	2.62
17	2.11	2.90	Sonsuz	1.96	2.58

*Bu tablo, *Biometrika for Statisticians*, vol. 1 (3rd ed.), New York: Cambridge University Press, 1970, edited by E. S. Pearson and H. O. Hartley'den, *Biometrika Trustees*'in izniyle uyarlanmıştır.

Bağımsız gruplarda *t* değerini hesaplama formülü:

$$t = \frac{\bar{X}_1 - \bar{X}_2}{s_{\bar{X}_1 - \bar{X}_2}}$$

bu formülde

$$s_{\bar{X}_1 - \bar{X}_2} = \sqrt{\left[\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}\right]\left[\frac{1}{n_1} + \frac{1}{n_2}\right]}$$

ve

$\bar{X}_1$ = Grup 1'in ortalaması
$\bar{X}_2$ = Grup 2'nin ortalaması
n_1 = Grup 1'in örneklem büyüklüğü
n_2 = Grup 2'nin örneklem büyüklüğü
s_1^2 = Grup 1'in değişkenliği
s_2^2 = Grup 2'nin değişkenliği
$N = n_1 + n_2$

TABLE A.3 *F*-DAĞILIMININ KRİTİK DEĞERLERİ*

Kullanım yönergesi: F'nin kritik değerini bulmak için tabloda, F-oranının paydasına ilişkin serbestlik derecesini içeren satır ile *F*-oranının payına ilişkin serbestlik derecesini içeren sütunun kesiştiği hücreyi bulunuz. Kalın yazı tipiyle listelenmiş sayılar, $\alpha = .05$ için olan kritik F değerleridir; kalın yazı tipiyle listelenmemiş olan sayılar ise, $\alpha = .01$ için olan kritik F değerleridir. Örnek vermek gerekirse, %5'lik anlamlılık düzeyini kabul ettiğimizi ve sdpay = 2 ve sdpayda = 12 serbestlik derecelerine sahip olan F değerinin anlamlılığını değerlendirdiğimizi varsayın. Tablodan, $\alpha = >05$ için kritik değeri, $F(2, 12) = 3.89$ olarak buluruz. Eğer elde edilen F değeri, bu kritik değeri geçerse, sıfır hipotezini reddederiz; elde edilen *F* değeri, bu kritik değerden küçükse, sıfır hipotezini reddetmeyiz.

Payda İçin Serbestlik Dereceleri	Pay İçin Serbestlik Dereceleri																	
	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	sonsuz
1	**161**	**200**	**216**	**225**	**230**	**234**	**237**	**239**	**241**	**242**	**244**	**246**	**248**	**249**	**250**	**251**	**252**	**254**
	4052	4999	5403	5625	5764	5859	5928	5981	6022	6056	6106	6157	6209	6325	6261	6287	6313	6366
2	**18.5**	**19.0**	**19.2**	**19.2**	**19.3**	**19.3**	**19.4**	**19.4**	**19.4**	**19.4**	**19.4**	**19.4**	**19.4**	**19.4**	**19.5**	**19.5**	**19.5**	**19.5**
	98.5	99.0	99.2	99.2	99.3	99.3	99.4	99.4	99.4	99.4	99.4	99.4	99.4	99.5	99.5	99.5	99.5	99.5
3	**10.1**	**9.55**	**9.28**	**9.12**	**9.01**	**8.94**	**8.89**	**8.85**	**8.81**	**8.79**	**8.74**	**8.70**	**8.66**	**8.64**	**8.62**	**8.59**	**8.57**	**8.53**
	34.1	30.8	29.5	28.7	28.2	27.9	27.7	27.5	27.4	27.2	27.0	26.9	26.7	26.6	26.5	26.4	26.3	26.1
4	**7.71**	**6.94**	**6.59**	**6.39**	**6.26**	**6.16**	**6.09**	**6.04**	**6.00**	**5.96**	**5.91**	**5.86**	**5.80**	**5.77**	**5.75**	**5.72**	**5.69**	**5.63**
	21.2	18.0	16.7	16.0	15.5	15.2	15.0	14.8	14.7	14.6	14.4	14.2	14.0	13.9	13.8	13.8	13.6	13.5
5	**6.61**	**5.79**	**5.41**	**5.19**	**5.05**	**4.95**	**4.88**	**4.82**	**4.77**	**4.74**	**4.68**	**4.62**	**4.56**	**4.53**	**4.50**	**4.46**	**4.43**	**4.26**
	16.3	13.3	12.1	11.4	11.0	10.7	10.5	10.3	10.2	10.0	9.89	9.72	9.55	9.47	9.38	9.29	9.20	9.02
6	**5.99**	**5.14**	**4.76**	**4.53**	**4.39**	**4.28**	**4.21**	**4.15**	**4.10**	**4.06**	**4.00**	**3.94**	**3.87**	**3.84**	**3.81**	**3.77**	**3.74**	**3.67**
	13.8	10.9	9.78	9.15	8.75	8.47	8.26	8.10	7.98	7.87	7.72	7.56	7.40	7.31	7.23	7.14	7.06	6.88
7	**5.59**	**4.74**	**4.35**	**4.12**	**3.97**	**3.87**	**3.79**	**3.73**	**3.68**	**3.64**	**3.57**	**3.51**	**3.44**	**3.41**	**3.38**	**3.34**	**3.30**	**3.23**
	12.2	9.55	8.45	7.85	7.46	7.19	6.99	6.84	6.72	6.62	6.47	6.31	6.16	6.07	5.99	5.91	5.82	5.65
8	**5.32**	**4.46**	**4.07**	**3.84**	**3.69**	**3.58**	**3.50**	**3.44**	**3.39**	**3.35**	**3.28**	**3.22**	**3.15**	**3.12**	**3.08**	**3.04**	**3.01**	**2.93**
	11.3	8.65	7.59	7.01	6.63	6.37	6.18	6.03	5.91	5.81	5.67	5.52	5.36	5.28	5.20	5.12	5.03	4.86
9	**5.12**	**4.26**	**3.86**	**3.63**	**3.48**	**3.37**	**3.29**	**3.23**	**3.18**	**3.14**	**3.07**	**3.01**	**2.94**	**2.90**	**2.86**	**2.83**	**2.79**	**2.71**
	10.6	8.02	6.99	6.42	6.06	5.80	5.61	5.47	5.35	5.26	5.11	4.96	4.81	4.73	4.65	4.57	4.48	4.31
10	**4.96**	**4.10**	**3.71**	**3.48**	**3.33**	**3.22**	**3.14**	**3.07**	**3.02**	**2.98**	**2.91**	**2.85**	**2.77**	**2.74**	**2.70**	**2.66**	**2.62**	**2.54**
	10.0	7.56	6.55	5.99	5.64	5.39	5.20	5.06	4.94	4.85	4.71	4.56	4.41	4.33	4.25	4.17	4.08	3.91
11	**4.84**	**3.98**	**3.59**	**3.36**	**3.20**	**3.09**	**3.01**	**2.95**	**2.90**	**2.85**	**2.79**	**2.72**	**2.65**	**2.61**	**2.57**	**2.53**	**2.49**	**2.40**
	9.65	7.21	6.22	5.67	5.32	5.07	4.89	4.74	4.63	4.54	4.40	4.25	4.10	4.02	3.94	3.86	3.78	3.60
12	**4.75**	**3.89**	**3.49**	**3.26**	**3.11**	**3.00**	**2.91**	**2.85**	**2.80**	**2.75**	**2.69**	**2.62**	**2.54**	**2.51**	**2.47**	**2.43**	**2.38**	**2.30**
	9.33	6.93	5.95	5.41	5.06	4.82	4.64	4.50	4.39	4.30	4.16	4.01	3.86	3.78	3.70	3.62	3.54	3.36
13	**4.67**	**3.81**	**3.41**	**3.18**	**3.03**	**2.92**	**2.83**	**2.77**	**2.71**	**2.67**	**2.60**	**2.53**	**2.46**	**2.42**	**2.38**	**2.34**	**2.30**	**2.21**
	9.07	6.70	5.74	5.21	4.86	4.62	4.44	4.30	4.19	4.10	3.96	3.82	3.66	3.59	3.51	3.43	3.34	3.17

14	**4.60**	**3.74**	**3.34**	**3.11**	**2.96**	**2.85**	**2.76**	**2.70**	**2.65**	**2.60**	**2.53**	**2.46**	**2.39**	**2.35**	**2.31**	**2.27**	**2.22**	**2.13**
	8.86	6.51	5.56	5.04	4.69	4.46	4.28	4.14	4.03	3.94	3.80	3.66	3.51	3.43	3.35	3.27	3.18	3.00
15	**4.54**	**3.68**	**3.29**	**3.06**	**2.90**	**2.79**	**2.71**	**2.64**	**2.59**	**2.54**	**2.48**	**2.40**	**2.33**	**2.29**	**2.25**	**2.20**	**2.16**	**2.07**
	8.68	6.36	5.42	4.89	4.56	4.32	4.14	4.00	3.89	3.80	3.67	3.52	3.37	3.29	3.21	3.13	3.05	2.87
16	**4.49**	**3.63**	**3.24**	**3.01**	**2.85**	**2.74**	**2.66**	**2.59**	**2.54**	**2.49**	**2.42**	**2.35**	**2.28**	**2.24**	**2.19**	**2.15**	**2.11**	**2.01**
	8.53	6.23	5.29	4.77	4.44	4.20	4.03	3.89	3.78	3.69	3.55	3.41	3.26	3.18	3.10	3.02	2.93	2.75
17	**4.45**	**3.59**	**3.20**	**2.96**	**2.81**	**2.70**	**2.61**	**2.55**	**2.49**	**2.45**	**2.38**	**2.31**	**2.23**	**2.19**	**2.15**	**2.10**	**2.06**	**1.96**
	8.40	6.11	5.18	4.67	4.34	4.10	3.93	3.79	3.68	3.59	3.46	3.31	3.16	3.08	3.00	2.92	2.83	2.65
18	**4.41**	**3.55**	**3.16**	**2.93**	**2.77**	**2.66**	**2.58**	**2.51**	**2.46**	**2.41**	**2.34**	**2.27**	**2.19**	**2.15**	**2.11**	**2.06**	**2.02**	**1.92**
	8.29	6.01	5.09	4.58	4.25	4.01	3.84	3.71	3.60	3.51	3.37	3.23	3.08	3.00	2.92	2.84	2.75	2.57
19	**4.38**	**3.52**	**3.13**	**2.90**	**2.74**	**2.63**	**2.54**	**2.48**	**2.42**	**2.38**	**2.31**	**2.23**	**2.16**	**2.11**	**2.07**	**2.03**	**1.98**	**1.88**
	8.18	5.93	5.01	4.50	4.17	3.94	3.77	3.63	3.52	3.43	3.30	3.15	3.00	2.92	2.84	2.76	2.67	2.49
20	**4.35**	**3.49**	**3.10**	**2.87**	**2.71**	**2.60**	**2.51**	**2.45**	**2.39**	**2.35**	**2.28**	**2.20**	**2.12**	**2.08**	**2.04**	**1.99**	**1.95**	**1.84**
	8.10	5.85	4.94	4.43	4.10	3.87	3.70	3.56	3.46	3.37	3.23	3.09	2.94	2.86	2.78	2.69	2.61	2.42
22	**4.30**	**3.44**	**3.05**	**2.82**	**2.66**	**2.55**	**2.46**	**2.40**	**2.34**	**2.30**	**2.23**	**2.15**	**2.07**	**2.03**	**1.98**	**1.94**	**1.89**	**1.78**
	7.95	5.72	4.82	4.31	3.99	3.76	3.59	3.45	3.35	3.26	3.12	2.98	2.83	2.75	2.67	2.58	2.50	2.31
24	**4.26**	**3.40**	**3.01**	**2.78**	**2.62**	**2.51**	**2.42**	**2.36**	**2.30**	**2.25**	**2.18**	**2.11**	**2.03**	**1.98**	**1.94**	**1.89**	**1.84**	**1.73**
	7.82	5.61	4.72	4.22	3.90	3.67	3.50	3.36	3.26	3.17	3.03	2.89	2.74	2.66	2.58	2.49	2.40	2.21
26	**4.23**	**3.37**	**2.98**	**2.74**	**2.59**	**2.47**	**2.39**	**2.32**	**2.27**	**2.22**	**2.15**	**2.07**	**1.99**	**1.95**	**1.90**	**1.85**	**1.80**	**1.69**
	7.72	5.53	4.64	4.14	3.82	3.59	3.42	3.29	3.18	3.09	2.96	2.81	2.66	2.58	2.50	2.42	2.33	2.13
28	**4.20**	**3.34**	**2.95**	**2.71**	**2.56**	**2.45**	**2.36**	**2.29**	**2.24**	**2.19**	**2.12**	**2.04**	**1.96**	**1.91**	**1.87**	**1.82**	**1.77**	**1.65**
	7.64	5.45	4.57	4.07	3.75	3.53	3.36	3.23	3.12	3.03	2.90	2.75	2.60	2.52	2.44	2.35	2.26	2.06
30	**4.17**	**3.32**	**2.92**	**2.69**	**2.53**	**2.42**	**2.33**	**2.27**	**2.21**	**2.16**	**2.09**	**2.01**	**1.93**	**1.89**	**1.84**	**1.79**	**1.74**	**1.62**
	7.56	5.39	4.51	4.02	3.70	3.47	3.30	3.17	3.07	2.98	2.84	2.70	2.55	2.47	2.39	2.30	2.21	2.01
40	**4.08**	**3.23**	**2.84**	**2.61**	**2.45**	**2.34**	**2.25**	**2.18**	**2.12**	**2.08**	**2.00**	**1.92**	**1.84**	**1.79**	**1.74**	**1.69**	**1.64**	**1.51**
	7.31	5.18	4.31	3.83	3.51	3.29	3.12	2.99	2.89	2.80	2.66	2.52	2.37	2.29	2.20	2.11	2.02	1.80
60	**4.00**	**3.15**	**2.76**	**2.53**	**2.37**	**2.25**	**2.17**	**2.10**	**2.04**	**1.99**	**1.92**	**1.84**	**1.75**	**1.7**	**1.65**	**1.59**	**1.53**	**1.39**
	7.06	4.98	4.13	3.65	3.34	3.12	2.95	2.82	2.72	2.63	2.50	2.35	2.20	2.12	2.03	1.94	1.84	1.60
120	**3.92**	**3.07**	**2.68**	**2.45**	**2.29**	**2.17**	**2.09**	**2.02**	**1.96**	**1.91**	**1.83**	**1.75**	**1.66**	**1.61**	**1.55**	**1.50**	**1.43**	**1.25**
	6.85	4.79	3.95	3.48	3.17	2.96	2.79	2.66	2.56	2.47	2.34	2.19	2.03	1.95	1.86	1.76	1.66	1.38
SONSUZ	**3.84**	**3.00**	**2.60**	**2.37**	**2.21**	**2.10**	**2.01**	**1.94**	**1.88**	**1.83**	**1.75**	**1.67**	**1.57**	**1.52**	**1.46**	**1.39**	**1.32**	**1.00**
	6.63	4.61	3.78	3.32	3.02	2.80	2.64	2.51	2.41	2.32	2.18	2.04	1.88	1.79	1.70	1.59	1.47	1.00

*Bu tablo, Biometrika Tables for Statisticians, cilt 1 (3. Baskı), New York: Cambridge University Press, 1970, editörler E. S. Pearson ve H. O. Hartley kaynağında yer alan Tablo 18'den, Biometrika Trustees'in izniyle özetlenmiştir.

SÖZLÜK

ABAB deseni-tersine desen (ABAB design–reversal design): Tek denekli deneysel desen; bu desende, başlangıçtaki temel düzey (A) belirlendikten sonra bağımsız değişken uygulanır (B); sonra bağımsız değişken kaldırılarak tekrar temel düzeye (A) dönülür; sonra tekrar bağımsız değişken uygulanır (B); deneyci, bağımsız değişkenin uygulanması ile davranışın değişip değişmediğini, kaldırılması ile bu değişikliğin ortadan kalkıp kalkmadığını gözler. Amaç, davranışta gözlenen değişmenin nedeninin, bağımsız değişken olup olmadığından emin olmaktır.

alan deneyi (field experiment): Gözlemcinin, davranış üzerindeki etkilerini belirlemek için bir veya daha fazla bağımsız değişkeni doğal bir ortamda değişimlemesi.

alfa (alpha): Bk. **anlamlılık düzeyi**.

alıştırma etkisi (practice effect): Birden fazla ölçüm alınması nedeniyle katılımcının uğradığı değişimler. Alıştırma etkisi, birçok defa ölçüm yapılmasına bağlı olarak ortaya çıkan hem pozitif (örneğin, bir göreve aşinalık) hem de negatif (örneğin, bıkkınlık) faktörlerin toplamıdır.

anket (questionnaire): Tarama tipi araştırmalarda bütün katılımcılar için başlıca veri toplama aracı olarak kullanılan önceden belirlenmiş soru seti.

anlamlılık düzeyi (level of significance): Sıfır hipotezi test edildiği zaman, bir sonucun istatistiksel olarak anlamlı olup olmadığına karar vermek için kullanılan olasılık değeri.

anlatımsal kayıt (narrative record): Bir davranışın, mümkün olduğunca aslına uygun bir kopyasını oluşturmak amacıyla kaydedilmesi.

ANOVA (ANOVA): Varyans analizi veya ANOVA, tek değişkenli bir araştırmada ya da birden fazla faktörlü (yani bağımsız değişkenli) araştırmalarda ikiden fazla ortalama karşılaştırılacağı zaman, sıfır hipotezini test etmek için en sık kullanılan çıkarımsal (inferential) testtir. ANOVA testi, bir deneydeki farklı varyans kaynaklarını analiz etmeye dayanır.

araç düzeni (instrumentation): Zamanla sadece katılımcılarda değil, aynı zamanda katılımcıların performansını ölçmek için kullanılan araçlarda da değişmeler meydana gelebilir. Araçların etkisine bağlı bu değişmeler, eğer işlemin (treatment) etkisinden ayrılamazsa, iç geçerliği tehdit edebilir.

aralık (range): Bir dağılımdaki en yüksek sayı ile en düşük sayı arasındaki fark. arasındaki kestirimsel ilişkileri belirlemek amacıyla yapılan araştırma.

araştırma sonrası bilgilendirme (debriefing): Verilerin toplandığı oturumun sonunda katılımcılara, katılmış oldukları araştırmanın amacı hakkında bilgi verme sürecidir. Katılımcılara, eğer yapılmışsa, deneyin amacı hakkında yanıltıcı bilgi verilmiş olmasının gerekçesi açıklanır ve yapmış oldukları katkının önemi vurgulanır. Bilgilendirmenin önemli amaçları, olası bazı yanlış anlamaları düzeltmek ve katılımcıların psikolojik araştırmalara dair olumlu izlenimlerle ayrılmalarını sağlamaktır.

araştırmada aldatma (deception): Bir araştırmanın amacı hakkında bazı bilgileri kasıtlı olarak katılımcıdan saklama veya yanlış bilgi verme.

ardışık bağımsız örneklemler deseni (successive independent samples design): Tarama tipi araştırma deseni; bir dizi kesitsel tarama yapılır ve sonraki her bir örneklemin üyelerine aynı sorular sorulur.

arşiv kayıtları (archival records): Bireylerin, kurumların, hükümetlerin ve diğer grupların faaliyetleriyle ilişkili kayıtlara veya belgelere dayanan kanıtlardan oluşan kaynaklar; diğer araştırma yöntemlerine alternatif olarak veya onlarla birlikte kullanılır.

aşırmacılık (plagiarism): Kaynağını açık şekilde belirtmeden, bir başkasının düşüncelerini veya çalışmasını sunma.

ayrımsal aktarma (differential transfer): Tekrarlı ölçümler deseninde bir koşuldaki performans kendisinden önceki koşuldan etkilendiği zaman gözlenen potansiyel bir problem.

bağımlı değişken (dependent variable): Araştırmacının, bağımsız değişkenin etkisini (eğer varsa) değerlendirmek için kullandığı davranış ölçüsü.

bağımsız değişken (independent variable): Araştırmacının, davranış üzerindeki etkisini belirlemek için en az iki düzeyini değişimlediği faktör.

bağımsız gruplar deseni (independent groups design): Deneydeki her bir bağımsız denek grubu, bağımsız değişkenin düzeyiyle belirlenen farklı bir koşulu temsil eder.

bağımsız gruplar *t* testi (t-test for independent groups): Farklı gruplardan alınan iki ortalamayı karşılaştırmak için kullanılan çıkarımsal bir test.

basit kesintili zaman-dizileri deseni (simple interrupted time series design): Bir yarı deneysel desendir; bu desende bir bağımlı değişkendeki değişmeler, bağımsız değişken uygulandıktan hem önce hem de sonra bir süre gözlenir.

basit seçkisiz örnekleme (simple random sampling): Bir olasılıklı örnekleme şekli; evren içinde belirlenmiş bir büyüklükteki mümkün her bir örneklemin seçilme olasılığı eşittir.

basit temel etki (simple main effect): Karmaşık bir desende bir bağımsız değişkenin, ikinci bir bağımsız değişkenin belirli bir düzeyindeki etkisi.

bilgilendirilmiş onam (informed consent): Araştırmanın niteliği, katılmamanın sonuçları ve katılma isteğini etkileyebilecek tüm faktörler hakkında yeterli bilgi verildikten sonra, araştırmaya katılma isteğini belirten açık ifade.

bilimsel yöntem (scientific method): Sezgisel bilgiden ziyade görgül bilgilere, test edilebilir hipotezlere, operasyonel olarak tanımlanmış olan olayların sistemli ve kontrollü gözlemine, doğru ve dakik araçlar kullanarak veri toplamaya, geçerli ve güvenilir ölçütlere ve sonuçların nesnel olarak rapor edilmesine odaklanan yaklaşım; bilim insanı eleştirel, daha da önemlisi, kuşkucu olmaya eğilimlidir.

bireysel farklar değişkeni (individual differences variable): Bireyler arasında sürekli olarak değişen özellik veya nitelik; depresyon düzeyi, yaş, zekâ, cinsiyet gibi. Bu değişken daha önceden (yani "doğal olarak") şekillendiği için, bir bireysel farklar değişkeni bazen doğal gruplar değişkeni olarak adlandırılır. Bazen bireysel farklar değişkeni ile eş anlamlı olarak kullanılan bir diğer terim, denek değişkeni'dir.

blok seçkisizleştirme (block randomization): Seçkisiz (random) gruplar deseninde seçkisiz atamayı sağlamak için en çok kullanılan teknik; her bir blok, koşulların seçkisiz şekilde sıralanmasından oluşur ve deneyin her bir koşulunda katılımcı sayısı kadar blok olur.

boylamsal desen (longitudinal design): Aynı katılımcılarla bir'den fazla görüşmenin yapıldığı araştırma.

çift körlemesine işlem (double-blind procedure): Hem katılımcı hem de araştırmacının hangi işlemin yapılmakta olduğunu bilmemesi.

Cohen *f* (Cohen's f): İkiden fazla ortalama olduğu zaman, ortalamalar arasındaki yayılımın derecesine göre hesaplanan bir etki büyüklüğü ölçüsü. f'nin .10, .25 ve .40 değerleri, sırasıyla, küçük, orta ve büyük etki olduğu anlamına gelir.

Cohen *d* (Cohen's d): Sık kullanılan bir etki büyüklüğü ölçümü; iki koşula ait ortalamaların farkı, katılımcıların puanlarının ortalama değişkenliğine (grup içi standart

sapma) bölünür. *d*'nin .20, .50 ve .80 değerleri sırasıyla, bir bağımsız değişkenin küçük, orta ve büyük etkisi olduğunu gösterir.

çok yöntemli yaklaşım (multimethod approach): Bir hipotezi test etmek amacıyla farklı araştırma yöntemleri ve davranış ölçümleri kullanarak veri toplayan yaklaşım; ölçme sürecinde davranışın bir tek gözlemi ile yetinmenin yetersiz olabileceği anlayışı.

çoklu temel düzey deseni–bireyler arası, davranışlar arası, durumlar arası (multiple baseline design–across individuals, across behaviors, across situations): Tek denekli bir deneysel desen; bu desende tedavinin etkili olduğuna karar vermek için, birden fazla temel düzeyi olan davranışın, tedavinin verilmesine bağlı olarak değiştiği gösterilir. Çoklu temel düzeyler farklı kişiler için, aynı kişinin farklı davranışları için, veya aynı kişi için farklı ortamlarda oluşturulur.

değişkenlik (variability): Bk. **dağılım ölçüleri**.

deney (experiment): Kontrollü bir araştırma yöntemi; araştırmacı bir veya daha fazla faktörü değişimler ve bu değişimlemenin davranışa olan etkisini gözler.

deneyci etkileri (experimenter effects): Deneyciyi, farklı araştırma gruplarındaki katılımcılara farklı şekilde davranmaya veya verileri kaydederken yanlı davranmaya sevk edebilen beklentileri.

denk olmayan kontrol grup deseni (nonequivalent control group design): Bir yarı deneysel araştırma deseni; bu desende karşılaştırma, katılımcıların gruplara seçkisiz yöntem dışındaki yöntemlerle atandığı kontrol ve deney grupları arasında yapılır.

denk olmayan kontrol gruplu zaman dizileri deseni (time series with nonequivalent control group design): Ayrıca Bk. **basit kesintili zaman dizileri deseni**. Bünyesine, denk olmayan bir kontrol grubu alarak basit zaman dizili bir desenin geçerliğini arttıran yarı deneysel işlem; uygulamadan hem önce hem de sonra, hem uygulama hem de kontrol grupları bir süre gözlenir.

dikkat çekmeyen–tepkisel olmayan ölçümler (unobtrusive–nonreactive–measures): Gözlenmekte olanlar gözlemcinin varlığını fark etmedikleri için, tepkisellik sorununu ortadan kaldıran ölçümler.

dış geçerlik (external validity): Bir araştırmanın sonuçlarının farklı evrenlere, ortamlara ve koşullara uygulanabilme derecesi.

doğal gözlem (naturalistic observation): Davranışın, gözlemcinin herhangi bir müdahalesinin olmadığı doğal bir ortamda gözlenmesi.

doğal gruplar deseni (natural groups design): Bir bağımsız gruplar deseni; bu desende koşullar, doğal olarak meydana gelen bağımsız bir değişkenin farklı düzeylerini temsil eder; örneğin, bireysel farklar değişkeni olarak yaş.

doğrusal eğilim (linear trend): Verilerin doğru bir çizgi hâlinde özetlenme eğilimi.

durum örnekleme (situation sampling): Koşullar, konumlar ve ortamlar arasında temsil edilirlik amacıyla durumların seçkisiz veya sistemli seçimi.

duyarlılık (sensitivity): Bir deneyde bağımsız değişkenin etkisinin, böyle bir etkisi gerçekten de var ise, ortaya çıkarılabilmesi olasılığıdır; hata değişkenliği azaldıkça (örneğin, değişkenleri dengelemekten ziyade, sabit tutarak) duyarlılık artar.

eşleştirilmiş gruplar deseni (matched groups design): Bağımsız gruplar deseninin bir şekli; araştırmacı, testten önce katılımcılara verdiği bir göreve dayanarak, katılımcıları denk gruplara ayırır; daha sonra, bu eşlenmiş grup üyelerini deneyin koşullarına seçkisiz olarak atar.

eta kare-η^2 (eta squared- η^2): Bağlantının (veya etki büyüklüğünün) kuvvetinin bir ölçüsü; bağımsız değişkenin bağımlı değişken üzerindeki etkisiyle açıklanan varyans oranına dayanır.

etki büyüklüğü (effect size): Örneklem büyüklüğünden bağımsız olarak, bağımsız değişken ile bağımlı değişken arasındaki ilişkinin kuvvetini belirten gösterge.

etkileşim (interaction effect): Bir bağımsız değişkenin etkisinin, ikinci bir bağımsız değişkenin düzeyine bağlı olarak değişmesi.

etnik merkezcilik (ethnocentrism): Sadece kendi kültüründeki deneyimlerine dayanarak farklı kültürlerdeki bireylerin davranışlarını açıklamaya çalışma.

evren (population): Hakkında araştırma yapılan veri kaynaklarının tamamı.

evrendeğer güven aralığı (confidence interval for a population parameter): Belirli (örneğin, %95) bir olasılıkla bir örneklem değer (örneğin, ortalama) çevresindeki değerler aralığı; evrendeğerin (örneğin, ortalamanın) bu aralığa düştüğü kabul edilir.

faktöryel desen (factorial design): Bk. **karmaşık desen**.

fiziksel izler (physical traces): Geçmişteki bir davranıştan arta kalanlara, onun parçalarına ve sonuçlarına dayanan veri kaynağı.

***F*-testi (*F*-test):** Varyans analizinde (ANOVA'da) gruplar arası değişkenlik ile grup içi veya hata değişkenliğinin oranı.

geçerlik (validity): Bir ölçümün doğruluğu; geçerli bir ölçüm, ölçtüğünü iddia ettiği şeyi ölçen ölçümdür.

genel *F*-testi (omnibus *F*-test): ANOVA'ya dayanılarak hesaplanan ilk, genel analiz.

gizlilik (privacy): Bireylerin, kendilerine ait bilgilerin başkalarına ne şekilde bildirileceğine karar verme hakkı.

görgül yaklaşım (empirical approach): Soruları cevaplamak için gerekli bilgileri ararken doğrudan gözlemi ve deney yapmayı önde tutan yaklaşım.

görüşmeci yanlılığı (interviewer bias): Görüşmeci bir sorunun cevabını kaynağa uydurmaya veya kaynağın cevaplarının bir kısmını seçerek kaydetmeye çalıştığı zaman ortaya çıkar.

gövde ve yaprak gösterimi (stem-and-leaf display): İlk basamaktaki sayılardan "gövdeler", virgülden sonraki sayılardan ise "yapraklar" oluşturarak, hem verilerin genel özelliklerini hem de spesifik item bilgilerini görselleştiren bir teknik.

gözlemci yanlılığı (observer bias): Sistemli gözlem hataları; çoğunlukla gözlemcinin bir araştırmanın sonuçlarına ilişkin beklentilerinden kaynaklanır (beklenti etkisi).

gözlemciler arası güvenirlik (interobserver reliability): İki bağımsız gözlemci arasındaki uyum derecesi.

güç (power): Bir istatistiksel testin, yanlış bir sıfır hipotezini reddetme olasılığı; güç, seçilen anlamlılık düzeyine, uygulamanın (treatment) etkisinin büyüklüğüne ve örneklemin büyüklüğüne bağlıdır.

güven aralığı (confidence interval): Belirli bir güven derecesiyle (örneğin, %95) evrene ait bir değeri kapsadığını düşündüğümüz değerler aralığını gösterir.

güvenirlik (reliability): Bir ölçüm, tutarlı olduğu zaman güvenilirdir.

Hawthorne etkisi (Howthorne effect): Bk. **yenilik etkileri**.

hipotez (hypothesis): Bir olayın geçici açıklaması.

iç geçerliğe yönelik tehditler (threats to internal validity): Açık bir neden-sonuç ilişkisi elde edebilmek için kontrol edilmesi gereken bir olayın olası nedenleri.

iç geçerlik (internal validity): Performanstaki farkların, kontrol edilemeyen bazı başka değişkenlere değil, açık bir şekilde bağımsız değişkenin etkisine bağlanabilme derecesi; iç geçerliği olan bir araştırma, bozucu faktörlerin etkisinin olmadığı bir araştırmadır.

içerik analizi (content analysis): Herhangi bir biçimde olabilmekle birlikte, çoğunlukla yazılı iletişimde mesajların niteliklerini nesnel şekilde belirleyerek çıkarımlar yapmak için kullanılan çeşitli tekniklerden herhangi birisi; arşiv verilerinin analizinde geniş şekilde kullanılır.

idiografik yaklaşım (idiographic approach): Hem birey olarak benzersizliği hem de meşruluğu önde tutarak, bir bireyin ayrıntılı şekilde incelenmesi.

iki ortalamanın karşılaştırılması (comparison of two means): Çoğunlukla, anlamlı çıkan bir *F* testinden sonra her defasında iki ortalamayı karşılaştırarak, bir deneydeki sistematik değişmenin spesifik kaynağını belirlemek için kullanılan bir istatistiksel tekniktir.

ilişkili bağımsız değişken (relevant independent variable): Davranışı, ya temel etki doğurarak doğrudan veya ikinci bağımsız değişken ile birlikte etkileşim doğurarak dolaylı yoldan etkilediği gösterilen bağımsız değişken.

istatistiksel olarak anlamlı (statistically significant): Bir deneyde bulunmuş olan farkların olasılığı beklenenden küçük olduğu zaman, eğer hata varyasyonunun tek başına bu farktan sorumlu olduğuna hükmediliyorsa, bulunan fark istatistiksel olarak anlamlıdır.

işevuruk tanım (operational definition): Bir kavramın sadece, ortaya çıkarılmasında ve ölçülmesinde kullanılan gözlenebilir işlemler açısından tanımlanması.

karıştırıcı (confounding): Etkisi araştırılmakta olan bağımsız değişken, sistemli bir şekilde, ikinci ve hesapta olmayan başka bir bağımsız değişkenle birlikte değişince meydana gelir.

karmaşık desen (complex design): İki veya daha fazla bağımsız değişkenin eş zamanlı olarak araştırıldığı deney.

karşıt dengeleme (counterbalancing): Tekrarlı ölçümler deseninde ortaya çıkabilen alıştırma etkisini farklı koşullara dengeli şekilde dağıtmaya yarayan kontrol tekniği. Karşıt dengelemenin başarısı, tamamlanmış tekrarlı ölçümler deseni mi yoksa tamamlanmamış tekrarlı ölçümler deseni mi kullanıldığına bağlıdır.

katılımcı gözlem (participant observation): Davranışın kaydedildiği ortamda veya bağlamda aynı zamanda aktif ve önemli bir rolü olan birisinin yaptığı davranış gözlemi.

katılımcı kaybı (subject attrition): İç geçerliğe bir tehdittir; örneğin, katılımcılar araştırma projesinden ayrıldıkları zaman meydana gelir. Katılımcı kaybı bir grubun niteliğini, bağımsız değişkenin uygulanmasından önce oluşturulmuş olan grubunkinden farklılaştırır; bunu örneğin, seçkisiz atamayla oluşturulmuş olan grup denkliğini bozarak yapar.

kayıp (attrition): Bk. **katılımcı kaybı.**

kesintili zaman dizileri deseni (interrupted time-series design): Bk. **basit kesintili zaman dizileri deseni** ve **denk olmayan kontrol gruplu zaman dizileri deseni.**

kesitsel desen (cross-sectional design): Tarama araştırması desenidir; evrenden bir veya daha fazla örneklem seçilir ve gerekli bilgiler bu örneklemlerden bir defada toplanır.

kirlenme (contamination): Katılımcı grupları deney hakkında birbirlerine bilgi verdikleri zaman ortaya çıkar.

kodlama (coding): Verilerin, özellikle de anlatıma dayanan verilerin işlenmesinde ilk aşama; kodlamada davranış veya özel olay birimleri belirli ölçütlere göre belirlenir ve sınıflanır.

kontrol (control): Bilimsel araştırmanın kilit ögesidir; bir olayın ortaya çıkmasında rolü olabilecek çeşitli faktörlerin etkileri izole edilir; üç temel kontrol şekli değişimleme (manipulation), koşulların sabit tutulması ve dengelemedir (balancing).

korelasyon (correlation): Aynı insanların, olayların veya şeylerin iki farklı ölçümünün birlikte değişmesidir; korelasyonun varlığı, bir değişkene ait değerler bilindiği zaman ikinci değişkene ait değerlerin kestirimini mümkün kılar.

korelasyon katsayısı (correlation coefficient): İki ölçümün birlikte değişme derecesini gösteren sayı; mutlak büyüklüğü .00 (ilişki yok) ile 1.00 (mükemmel ilişki) arasında değişir; birlikte değişimin yönü, katsayının işaretiyle belirtilir: Artı (+) işareti iki değişkenin aynı yönde birlikte değiştiğini, eksi (-) işareti ise iki değişkenin zıt yönlerde değiştiğini gösterir.

korelasyonel araştırma (correlational research): Doğal olarak meydana gelen değişkenler

küçük n araştırması (small-n research): Bk. **tek denekli deney**.

mekanik denek kaybı (mechanical subject loss): Araç gereç sorunu yüzünden veya deneycinin hatası yüzünden katılımcının deneyi tamamlayamaması.

merkezî eğilim (central tendency): Bk. **merkezî eğilim ölçüleri**.

merkezî eğilim ölçüleri (measures of central tendency): Ortalama, ortanca ve mod gibi ortalama etrafındaki bir puanı gösteren ölçümler.

meta-analiz (meta-analysis): Aynı konuda yapılmış olan çok sayıdaki bağımsız araştırmanın sonuçlarına uygulanan analiz; meta-analizde kullanılan tipik ölçüm, etki büyüklüğüdür.

minimal risk (minimal risk): Katılımcının araştırmada maruz kalması beklenen zarar veya rahatsızlığın olasılık ve büyüklüğü, günlük hayatta veya olağan bir ölçüm sırasında karşılaşılandan daha büyük olmadığı zaman, bir katılımcının minimal risk altında kaldığı düşünülür.

mod (mode): Dağılımda en sık rastlanan puan.

***N* = 1 deseni (*N* = 1 design):** Bk. **tek denekli desen**.

nedensel çıkarım (causal inference): Neden ve sonucun birlikte değiştiğini, etkiden önce düzenli şekilde nedenin geldiğini belirleyerek ve olası alternatif nedenleri eleyerek bir olayın nedenini veya nedenlerini belirleme.

negatif korelasyon (negative correlation): İki değişken arasındaki ilişki; değişkenlerden birisinin değeri düştükçe, diğerinin değeri yükselir.

nomotetik yaklaşım (nomothetic approach): Büyük gruplar (evrenler) için geçerli olacak geniş genellemelere veya yasalara ulaşmayı amaçlayan yaklaşım; bir grubun ortalamasına veya tipik performansına ağırlık verilir.

olasılığa dayanan örnekleme (probability sampling): Evrenin her bir elemanının örnekleme alınması olasılığının belirlendiği örneklem seçme işlemi.

olasılığa dayanmayan örnekleme (nonprobability sampling): Üyelerin örnekleme alınması olasılığının tahmin edilemediği bir örneklem belirleme işlemi; yaygın bir şekli, uygun örnekleme'dir.

ölçek türü (measurement scale): Fiziksel ve psikolojik dört ölçüm düzeyinin her birisi: Kategorik (nominal), sıralı (ordinal), eşit aralıklı (interval) ve oranlı (ratio) ölçek. Eşit aralıklı ölçekte uyarımlar arasındaki uzaklık belirtilir; oranlı ölçekte ise gerçek bir sıfır noktası vardır.

olgunlaşma (maturation): Sadece zamanın geçmesine bağlı olarak ortaya çıkan değişme; bir deneyde katılımcıların, deneycinin uyguladığı işlemler nedeniyle değil de olgunlaşmaya bağlı olarak uğradıkları değişmeler, iç geçerliği zayıflatabilir.

örneklem (sample): Araştırmaya konu olan ögelerin tamamından az olan şey; tarama araştırmalarında, gerçekte araştırılmakta olan evrenin bir alt grubu.

ortalama (mean): Aritmetik ortalama, puanların toplamı, o toplama dahil olan puanların sayısına bölünerek elde edilir. Ortalama, en sık kullanılan merkezî eğilim ölçüsüdür.

ortalamanın standart hatası (standard error of the mean): Ortalamaların örnekleme dağılımının standart sapması.

ortalamanın tahminî standart hatası (estimated standard error of the mean): Örneklemin standart sapmasının örneklem büyüklüğünün kare köküne bölünmesiyle elde edilen gerçek standart hatanın bir tahmini.

ortanca (median): Bir dağılımın orta noktası; dağılımdaki puanların yarısı bu noktanın üzerinde, diğer yarısı ise altında kalır.

plasebo kontrol grubu (placebo control group): Bir ilaca veya başka bir aktif maddeye benzemekle birlikte gerçekte etkisiz (aktif olmayan) bir maddenin katılımcılara verilmesi işlemi.

pozitif korelasyon (positive correlation): İki değişken arasındaki ilişki; değişkenlerden birisinin değeri arttıkça (veya azaldıkça), diğerinin değeri de yükselir (veya azalır).

program değerlendirme (program evaluation): Bir kurumun, bir devlet görevlisinin veya toplumun başka bir biriminin önerdiği bir değişikliğin gerekli olup olmadığını ve planlanan etkiyi yapıp yapmayacağını ve tamamlandığı zaman, makul bir maliyetle, arzu edilen sonucu doğurup doğurmayacağını belirleme.

regresyon–ortalamaya doğru (regression–to the mean): Bireyler bir deneye "aşırı" puanları nedeniyle katılmak üzere seçildikleri zaman, istatistiksel gerileme meydana gelebilir. İstatistiksel gerileme iç geçerliğe bir tehdittir, çünkü aşırı gruplardan seçilmiş olan bireylerin, *herhangi bir işleme tâbi tutulmadan, sadece istatistiksel gerileme nedeniyle* ikinci bir testte ("son test") daha az aşırı puanlar almaları beklenir.

risk/fayda oranı (risk/benefit ratio): Önerilen araştırmanın sonuçlarının katılımcıya ve topluma kazandıracaklarına kıyasla, katılımcıyı maruz bıraktığı riskin öznel değerlendirmesi.

saçılım grafiği (scatterplot): Aynı kişi, şey veya olaydan alınmış olan iki ölçümün kesişme noktasını işaret ederek, iki değişken arasındaki ilişkiyi gösteren grafik.

sahte ilişki (spurious relationship): Bulguların, yanlış şekilde, iki veya daha fazla değişken arasında ilişki olduğunu göstermesi.

seçici birikim (selective deposit): Fiziksel kanıtların saklanma şeklinden, arşiv kaynaklarının oluşturulma düzenlenme ve değiştirilme şeklinden kaynaklanan yanlılık; bu yanlılığın varlığı, araştırma bulgularının genellenmesini büyük ölçüde engeller.

seçici denek kaybı (selective subjects loss): Katılımcılar, araştırmanın sonuçlarıyla ilişkili bazı özellikleri nedeniyle, araştırma boyunca farklı gruplarda farklı oranlarda kaybolduğu zaman ortaya çıkar.

seçici kalıcılık (selective survival): Fiziksel kanıtların ve arşiv kayıtlarının korunma şeklinden kaynaklanan yanlılık; bu yanlılığın olması, araştırma bulgularının dış geçerliğini önemli ölçüde sınırlar.

seçim (selection): Seçme, iç geçerliğe bir tehdittir; araştırmanın başında, deneyin bir grubundaki bireylerin özellikleri ile diğer grubundaki bireylerin özellikleri arasında farklar olduğu zaman ortaya çıkar.

seçim yanlılığı (selection bias): Örneklemin temsil ediciliğine tehdit; bir örneklem seçmek için kullanılan işlemler evrenin önemli bir kısmının, olması gerekenden daha fazla veya daha az temsil edilmesine yol açtığı zaman ortaya çıkar.

seçkisiz atama (random assignment): Bir bağımsız gruplar deseninde grupları oluşturmak için en çok kullanılan teknik; amaç, bireysel farklılıkları dengeleyerek, denk gruplar kurmaktır.

seçkisiz gruplar deseni (random groups design): Bağımsız gruplar deseninin en yaygın şekli; bu desende katılımcılar gruplara seçkisiz olarak atanır; böylece, deneyin başında grupların denk olduğu kabul edilir.

seçkisiz örnekleme (random sampling): Bk. **basit seçkisiz örnekleme**

sıfır hipotezi anlamlılık testi SHAT (null hypothesis significance testing NHST): Bir araştırmada bir değişkenin etkisinin olup olmadığına karar vermek için kullanılan bir istatistiksel çıkarım işlemidir. Bu işlem, söz konusu değişkenin bir etkisinin olmadığı (bk. **sıfır hipotezi**) varsayımı ile yola çıkar ve bir araştırmada gözlenen etkinin (örneğin, koşullar arasındaki ortalama farkının) sadece hata değişkenliğinden ("tesadüf") kaynaklanmış olması olasılığını belirlemek için olasılık teorisi kullanılır. Sıfır hipotezinin doğru olduğu varsayıldığında, böyle bir etkinin ortaya çıkma olasılığı küçük ise (bk. **anlamlılık düzeyi**), değişkenin güvenilir bir etki doğurduğu sonucuna varırız (bk. **istatistiksel olarak anlamlı**)

sıfır hipotezi H_0 (null hypothesis H_0): İstatistiksel çıkarımda ilk adım olarak kullanılan varsayım; buna göre, bağımsız değişkenin bir etkisi yoktur.

sınama etkisi (testing): Bir teste tâbi tutulma, genellikle bir sonraki testin performansını etkiler. Eğer bağımsız değişkenin etkisi test işleminin bu etkisinden ayrılamazsa, test etkisi iç geçerliği tehdit edebilir.

sosyal beğenirlik (social desirability): Katılımcı üzerindeki baskı; sosyal açıdan en kabul edilebilir şekilde cevap verilmesi gerektiğini düşünerek cevap verme eğilimi.

standart sapma (standard deviation): En çok kullanılan dağılım ölçüsü; puanların, ortalama olarak, ortalamadan ne kadar farklı olduğunu gösterir.

tabakalı seçkisiz örnekleme (stratified random sampling): Bir olasılıklı örnekleme tekniği; evren, "tabaka" adı verilen alt evrenlere bölünür ve örneklemler bu tabakaların her birisinin içinden seçkisiz olarak alınır.

taban etkisi (floor effect): Bk. **tavan etkisi**.

talep özellikleri (demand characteristics): Bir araştırmada katılımcının ne yapması gerektiğine karar verirken kullandığı ipuçları ve diğer bilgiler; bu ipuçları ve bilgiler çoğunlukla katılımcıyı, araştırmacının beklentisi olduğuna inandığı şeyleri yapmaya sevk eder.

tarih etkisi (history): Bağımsız değişken dışında, katılımcıların davranışında değişikliğe yol açtığı takdirde iç geçerliği tehdit edebilen bir olayın meydana gelmesi.

tavan ve taban etkisi (ceiling and floor effect): Deneyin herhangi bir koşulunda katılımcıların performansı en üst (veya en alt) seviyeye ulaştığı için, deneycinin bir bağımsız değişkenin veya mümkün bir etkileşimin etkisini ölçememesi problemi.

tek denekli deney (single subject experiment): Bir tek kişide, davranışı sürekli olarak gözlerken, koşulları sistemli şekilde karşılaştırarak davranıştaki değişmelere odaklanan işlem.

tek faktörlü bağımsız gruplar deseni (single-factor independent groups design): Tek bağımsız değişkenin ve bağımsız grupların kullanıldığı bir deney şekli.

tekrar çalışması (replication) Aynı sonuçların elde edilip edilemeyeceğini belirlemek için, daha önce yapılmış olan bir deneydeki işlemi aynen kullanarak yapılan çalışma.

tekrarlı ölçümler (denek içi) *t*-testi (repeated measures–within-subjects–*t*-test): Aynı grup üyelerine ait veya bağımlı değişkenle ilişkili bazı kriterler açısından "eşleştirilmiş" iki ayrı grup katılımcıya ait iki ortalamayı karşılaştırmaya yarayan çıkarımsal bir test.

tekrarlı ölçümler desenleri (repeated measures designs): Her bir katılımcının, deneyin bütün koşullarına katıldığı araştırma desenleri (yani, aynı katılımcıdan birden fazla ölçüm alınır).

temel araştırma (basic research): Davranış ve zihinsel süreçler hakkındaki anlayışımızı geliştirecek bilgilere ulaşmayı amaçlayan araştırmalar. Ayrıca bk. **uygulamalı araştırma**.

temel düzey aşaması (baseline stage): Tek denekli deneysel desenin ilk aşamasıdır; bu aşamada, davranışın herhangi bir müdahale yapılmadan önceki düzeyi ölçülür.

temel etki (main effect): Karmaşık bir araştırma deseninde bir bağımsız değişkenin toplam etkisi.

temsil gücü (representativeness): Bir örneklemin özellikleri ile içinden alındığı evrenin özellikleri örtüştüğü ölçüde, o örneklem temsil edici olur; örneklemden evrene genelleme yapabilme gücümüzü en çok etkileyen faktör, bu temsil edicilik düzeyidir.

teori (theory): Olayları belirlemek, olaylar arasındaki ilişkileri betimlemek ve bu olayların nasıl meydana geldiklerini açıklamak için öne sürülmüş, mantıksal olarak organize önermeler.

tepki oranı yanlılığı (response rate bias): Bir örneklemin temsil ediciliğine tehdit; seçilen katılımcılardan bazıları anketi tamamlamada sistemli şekilde başarısız oldukları (örneğin, katılımcı çok uzun bir anketi tamamlayamadığı veya telefonla yapılan bir ankette kendisinden istenilene uymadığı) zaman meydana gelir.

tepkisellik (reactivity): Bir gözlemcinin, gözlenmekte olan davranış üzerindeki etkisi; gözlemcinin etkilediği davranış, bir gözlemcinin olmadığı koşulda meydana gelen davranışı temsil etmeyebilir.

tersine desen (reversal design): Bk. **ABAB deseni**.

Tip I hata (Type I error): Sıfır hipotezini, doğru olduğu zaman, reddetme olasılığı; anlamlılık veya alfa düzeyine eşittir.

Tip II hata (Type II error): Sıfır hipotezini, yanlış olduğu zaman, kabul etme olasılığı.

uygulamalı araştırma (applied research): Bir durumu geliştirmeye yarayacak nitelikte bilgi elde etmeyi amaçlayan araştırma. Ayrıca bk. **Temel araştırma**.

vaka çalışması (case study): Sadece bir kişinin belirli bir açıdan ayrıntılı şekilde betimlenmesi ve analizi.

veri analiz aşamaları (stages of data analysis): Veri analizinin üç aşaması; verileri tanıma, verileri özetleme ve verilerin ortaya ne çıkardığından emin olmadır.

verileri özetleme (summarizing the data): Veri analizinin bu ikinci aşamasında araştırmacı, verilerin içerdiği bilgileri özetlemek için betimsel istatistik ve grafikler kullanır. Verilerdeki eğilimler ve örüntüler betimlenir.

verileri tanımaya başlamak (getting to know the data): Verilerin analizinin bu ilk aşamasında araştırmacı verilerdeki hataları, uç değerleri denetler ve verilerin genel özellikleri hakkında bir kanaat edinir.

verilerin anlamının onaylanması (confirming what the data reveal): Verilerin analizinin üçüncü aşamasında araştırmacı, verilerin davranış hakkında ne söylediğini belirler. istatistiksel teknikler, sonuçların "tesadüfi" olduğunu öne süren iddiaların yanlış olduğunu göstermek için kullanılır.

verilerin indirgenmesi (data reduction): Verilerin analizi sürecinin bu aşamasında sonuçlar anlamlı bir şekilde organize edilir ve önemli bulguları özetleyen ifadeler hazırlanır.

yapı (construct): Davranışı veya zihinsel süreçleri açıklamak için psikoloji teorilerinin kullandıkları kavram veya düşünce; örneğin, saldırganlık, depresyon, zekâ, bellek, ve kişilik.

yapılandırılmış gözlem (structured observation): Müdahale kullanan bir gözlem yöntemi; kontrol derecesi çoğunlukla alan araştırmalarına kıyasla daha azdır; sıklıkla, davranış değerlendirmeleri yapacakları zaman, klinik ve gelişim psikologları tarafından kullanılır.

yarı deneysel desenler (quasy-experiments): Gerçek deneylere benzer özellikleri olan işlemler; örneğin, bazı tip müdahaleler veya uygulamalar kullanılır ve bir karşılaştırma yapılır ama, gerçek bir deneyde olduğu gibi, farklı kontrol dereceleri yoktur.

yayılım ölçüleri–değişkenlik (measures of dispersion–varibiality): Aralık (range), standart sapma gibi bir dağılımdaki sayıların dağılımını betimleyen ölçümler.

yenilik etkileri (novelty effects): Araştırmanın iç geçerliğine yönelik bir tehdit; insanların davranışları sadece bir yeniliğin (örneğin, bir tedavi) yol açtığı heyecan, enerji, ve coşku nedeniyle değiştiği zaman ortaya çıkar; Hawthorne etkisi, yenilik etkisinin özel bir şeklidir.

zaman örneklemesi (time sampling): Davranışın temsil edici bir örneklemini elde etmek için gözlem aralıklarının sistemli veya seçkisiz bir şekilde seçilmesi.

Kaynakça

Abelson, R. P. (1995). *Statistics as principled argument*. Hillsdale, NJ: Erlbaum.

Abelson, R. P. (1997). On the surprising longevity of flogged horses: Why there is a case for the significance test. *Psychological Science, 8*, 12–15.

Adler, T. (1991, December). Outright fraud rare, but not poor science. *APA Monitor*, 11.

Allport, G. W. (1961). *Pattern in growth and personality*. New York: Holt, Rinehart and Winston.

Ambady, N., & Rosenthal, R. (1993). Half a minute: Predicting teacher evaluations from thin slices of nonverbal behavior and physical attractiveness. *Journal of Personality and Social Psychology, 64*, 431–441.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Washington, DC: Author.

American Psychological Association. (2002). Ethical principles of psychologists and code of conduct. *American Psychologist, 57*, 1060–1073.

American Psychological Association. (2010a). 2010 Amendments to the 2002 "Ethical principles of psychologists and code of conduct." *American Psychologist, 65*(5), 493.

American Psychological Association. (2010b). *Publication Manual* (6th ed.). Washington, DC: Author.

American Psychological Association. (2012). Summary report of journal operations, 2011. *American Psychologist, 67*(5), 410–411.

Anand, V. (2007). A study of time management: The correlation between video game usage and academic performance markers. *CyberPsychology & Behavior, 10*, 552–559. doi: 10.1089/cpb.2007.9991

Anderson, C. A., Berkowitz, L., Donnerstein, E., Huesmann, L. R., Johnson, J. D., Linz, D., Malamuth, N. M., & Wartella, E. (2003). The influence of media violence on youth. *Psychological Science in the Public Interest, 4*, 81–110.

Anderson, C. A., & Bushman, B. J. (1997). External validity of "trivial" experiments: The case of laboratory aggression. *Review of General Psychology, 1*, 19–41.

Anderson, C. R. (1976). Coping behaviors as intervening mechanisms in the inverted-U stress-performance relationship. *Journal of Applied Psychology, 61*, 30–34.

Anderson, J. R. (1990). *The adaptive character of thought*. Hillsdale, NJ: Erlbaum.

Anderson, K. J., & Revelle, W. (1982). Impulsivity, caffeine, and proofreading: A test of the Easterbrook hypothesis. *Journal of Experimental Psychology: Human Perception and Performance, 8*, 614–624.

Anglin, J. M. (1993). Vocabulary development: A morphological analysis. *Monographs of the Society for Research in Child Development, 58* (10, Serial No. 238).

Arkowitz, H., & Lilienfeld, S. O. (2006, October/November). Do self-help books help? *Scientific American Mind, 17*, 90–91. doi: 10.1038/scientificamericanmind0406-42/

Armstrong-Stassen, M. (1994). Coping with transition: A study of layoff survivors. *Journal of Organizational Behavior, 15*, 597–621. doi: 10.1002/job.4030150705

Armstrong-Stassen, M. (2006). Determinants of how managers cope with organizational downsizing. *Applied Psychology: An International Review, 55*, 1–26. doi: 10.1111/j.1464-0597.2006.00225.x

Arnett, J. J. (2008). The neglected 95%: Why American psychology needs to become less American. *American Psychologist, 63*, 602–614. doi: 10.1037/0003-066x.63.7.602

Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In K. W. Spence & J. T. Spence (Eds.), *The psychology of learning and motivation* (Vol. 2, pp. 89–195). New York: Academic Press.

Back, M. D., Kufner, A. C. P., & Egloff, B. (2010). The emotional timeline of September 11, 2001. *Psychological Science, 21,* 1417–1419. doi: 10.1177/0956797610382124

Back, M. D., Stopfer, J. M., Vazire, S., Gaddis, S., Schmukle, S. C., Egloff, B., & Gosling, S. (2010). Facebook profiles reflect actual personality, not self-idealization. *Psychological Science, 21,* 372–374. doi: 10.1177/0956797609360756

Baicker, K., Taubman, S. L., Allen, H. L., Bernstein, M., Gruber, J. H., Newhouse, J. P., . . . Finkelstein, A. N. (2013). The Oregon experiment—Effects of Medicaid on clinical outcomes. *The New England Journal of Medicine, 368,* 1713–1722. doi: 10.1056/NEJMsa1212321

Baker, T. B., McFall, R. M., & Shoham, V. (2008). Current status and future prospects of clinical psychology: Toward a scientifically principled approach to mental and behavioral health care. *Psychological Science in the Public Interest, 9,* 67–103. doi: 10.1111/j.1539-6053.2009.01036.x

Banaji, M. R., & Crowder, R. G. (1989). The bankruptcy of everyday memory. *American Psychologist, 44,* 1185–1193.

Bard, K. A., Myowa-Yamakoshi, M., Tomonaga, M., Tanaka, M., Costall, A., & Matsuzawa, T. (2005). Group differences in the mutual gaze of chimpanzees (*Pan Troglodytes*). *Developmental Psychology, 41,* 616–624.

Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology, 51,* 1173–1182.

Bartholomew, G. A. (1982). Scientific innovation and creativity: A zoologist's point of view. *American Zoologist, 22,* 227–335.

Bartlett, M. Y., & DeSteno, D. (2006). Gratitude and prosocial behavior: Helping when it costs you. *Psychological Science, 17,* 319–325.

Bathje, G. J., & Pryor, J. B. (2011). The relationships of public and self-stigma to seeking mental health services. *Journal of Mental Health Counseling, 33*(2), 161–176.

Baumeister, R. F., Vohs, K. D., & Funder, D. C. (2007). Psychology as the science of self-reports and finger movements: Whatever happened to actual behavior? *Perspectives on Psychological Science, 2,* 396–403. doi: 10.1111/j.1745-6916.2007.00051.x

Baumrind, D. (1985). Research using intentional deception: Ethical issues revisited. *American Psychologist, 40,* 165–174.

Becker-Blease, K. A., & Freyd, J. J. (2006). Research participants telling the truth about their lives: The ethics of asking and not asking about abuse. *American Psychologist, 61,* 218–226.

Behnke, S. (2003). Academic and clinical training under APA's new ethics code. *Monitor on Psychology, 34,* 64.

Behnke, S. (2009). A classic study revisited. *Monitor on Psychology, 40*(4), 76.

Berk, R. A., Boruch, R. F., Chambers, D. L., Rossi, P. H., & Witte, A. D. (1987). Social policy experimentation: A position paper. In D. S. Cordray & M. W. Lipsey (Eds.), *Evaluation Studies Review Annual* (Vol. 11, pp. 630–672). Newbury Park, CA: Sage.

Birnbaum, M. H. (2000). Decision making in the lab and on the Web. In M. H. Birnbaum (Ed.), *Psychological experiments on the Internet* (pp. 3–34). San Diego, CA: Academic Press.

Birnbaum, M. (2004). Human research and data collection. *Annual Review of Psychology, 55,* 803–832. doi: 10.1146/annurev.psych.55.090902.141601

Blanchard, F. A., Crandall, C. S., Brigham, J. C., & Vaughn, L. A. (1994). Condemning and condoning racism: A social context approach to interracial settings. *Journal of Applied Psychology, 79,* 993–997.

Blanck, P. D., Bellack, A. S., Rosnow, R. L., Rotheram-Borus, M. J., & Schooler, N. R. (1992). Scientific rewards and conflicts of ethical choices in human subjects research. *American Psychologist, 47*, 959–965.

Blass, T. (2009). From New Haven to Santa Clara: A historical perspective on the Milgram obedience experiments. *American Psychologist, 64*, 37–45. doi: 10.1037/a0014434

Bolgar, H. (1965). The case study method. In B. B. Wolman (Ed.), *Handbook of clinical psychology* (pp. 28–39). New York: McGraw-Hill.

Bonanno, G. A. (2004). Loss, trauma, and human reliance: Have we underestimated the human capacity to thrive after extremely aversive events? *American Psychologist, 59*, 20–28. doi: 10.1037/0003-066X.59.1.20

Boring, E. G. (1954). The nature and history of experimental control. *American Journal of Psychology, 67*, 573–589.

Bowker, A., Boekhoven, B., Nolan, A., Bauhaus, S., Glover, P., Powell, T., & Taylor, S. (2009). Naturalistic observations of spectator behavior at youth hockey games. *The Sport Psychologist, 23*(3), 301–316.

Brandt, R. M. (1972). *Studying behavior in natural settings.* New York: Holt, Rinehart and Winston: University Press of America, 1981.

Brotsky, S. R., & Giles, D. (2007). In the "Pro-ana" community: A covert online participant observation. *Eating Disorders: The Journal of Treatment & Prevention, 15*(2), 93–109.

Brown, R., & Kulik, J. (1977). Flashbulb memories. *Cognition, 5*, 73–99.

Buchanan, T. (2000). Potential of the Internet for personality research. In M. H. Birnbaum (Ed.), *Psychological experiments on the Internet* (pp. 121–139). San Diego, CA: Academic Press.

Buchanan, T., & Williams, J. E. (2010). Ethical issues in psychological research on the Internet. In S. D. Gosling & J. A. Johnson (Eds.), *Advanced methods for conducting online behavioral research* (pp. 255–271). Washington, DC: American Psychological Association.

Buhrmester, M., Kwang, T., & Gosling, S. D. (2011). Amazon's Mechanical Turk: A new source of inexpensive, yet high-quality data? *Perspectives on Psychological Science, 6*, 3–5. doi: 10.1177/1745691610393980

Burger, J. M. (2007). Replicating Milgram. *APS Observer, 20*(11), 15–17.

Burger, J. M. (2009). Replicating Milgram: Would people still obey today? *American Psychologist, 64*, 1–11. doi: 10.1037/a0010932

Bushman, B. J. (2005). Violence and sex in television programs do not sell products in advertisements. *Psychological Science, 16*, 702–708.

Campbell, D. T. (1969). Reforms as experiments. *American Psychologist, 24*, 409–429.

Campbell, D. T., & Stanley, J. C. (1966). *Experimental and quasi-experimental designs for research.* Chicago: Rand McNally.

Candland, D. K. (1993). *Feral children and clever animals.* New York: Oxford University Press.

Carnagey, N. L., & Anderson, C. A. (2005). The effects of reward and punishment in violent video games on aggressive affect, cognition, and behavior. *Psychological Science, 16*, 882–889.

Ceci, S. J. (1993). Cognitive and social factors in children's testimony. Master lecture presented at the American Psychological Association Convention.

Chabris, C. F. (1999). Prelude or requiem for the "Mozart effect"? *Nature, 400*, 826–827. doi: 10.1038/23608

Chabris, C. F., & Simons, D. (2010). *The invisible gorilla: How our intuitions deceive us.* New York, NY: Crown.

Chastain, G., & Landrum, R. E. (1999). *Protecting human subjects: Departmental subject pools and institutional review boards.* Washington, DC: American Psychological Association.

Chernoff, N. N. (2002, December). Nobel Prize winner pushes economic theory despite hurdles. *APS Observer, 15*, 9–10.

Chow, S. L. (1988). Significance test or effect size? *Psychological Bulletin, 103*, 105–110.

Christensen, L. (1988). Deception in psychological research: When is its use justified? *Personality and Social Psychology Bulletin, 14*, 664–675.

Cohen, A. B. (2009). Many forms of culture. *American Psychologist, 64*, 94–104. doi: 10.1037/a0015308

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Cohen, J. (1990). Things I have learned (so far). *American Psychologist, 45*, 1304–1312.

Cohen, J. (1992). A power primer. *Psychological Bulletin, 112*, 155–159.

Cohen, J. (1995). The earth is round ($p < .05$). *American Psychologist, 49*, 997–1003.

Congressional Budget Office. (2012, January). *Lessons from Medicare's demonstration projects on disease management, care coordination, and value-based payment* (Issue Brief). Washington, DC: Author.

Conway, A. R. A., Skitka, L. J., Hemmerich, J. A., & Kershaw, T. C. (2009). Flashbulb memory for 11 September 2001. *Applied Cognitive Psychology, 23*, 605–623. doi: 10.1002/acp.1497

Cook, T. D., & Campbell, D. T. (1979). *Quasi-experimentation: Design and analysis issues for field settings*. Chicago: Rand McNally.

Coon, D. J. (1992). Testing the limits of sense and science: American experimental psychologists combat spiritualism, 1880–1920. *American Psychologist, 47*, 143–151.

Cordaro, L., & Ison, J. R. (1963). Psychology of the scientist: X. Observer bias in classical conditioning of the planarian. *Psychological Reports, 13*, 787–789.

Corrigan, P. (2004). How stigma interferes with mental health care. *American Psychologist, 59*, 614–625. doi: 10.1037/0003-066X.59.7.614

Creswell, J. W. (2013). *Qualitative inquiry and research design: Choosing among five approaches* (3rd ed.). Thousand Oaks, CA: Sage.

Crick, N. R., Ostrov, J. M., Burr, J. E., Cullerton-Sen, C., Jansen-Yeh, E., & Ralston, P. (2006). A longitudinal study of relational and physical aggression in preschool. *Applied Developmental Psychology, 27*(3), 254–268.

Cronbach, L. J. (1992). Four *Psychological Bulletin* articles in perspective. *Psychological Bulletin, 12*, 389–392.

Crossen, C. (1994). *Tainted truth: The manipulation of fact in America*. New York: Simon & Schuster.

Cumming, G., & Finch, S. (2005). Inference by eye: Confidence intervals and how to read pictures of data. *American Psychologist, 60*, 170–180.

Curtiss, S. R. (1977). *Genie: A psycholinguistic study of a modern-day "wild child."* New York: Academic Press.

Dallam, S. J., Gleaves, D. H., Cepeda-Benito, A., Silberg, J. L., Kraemer, H. C., & Spiegel, D. (2001). The effects of child sexual abuse: Comment on Rind, Tromovitch, and Bauserman (1998). *Psychological Bulletin, 127*, 715–733.

Dawes, R. M. (1991, June). *Problems with a psychology of college sophomores*. Paper presented at the Third Annual Convention of the American Psychological Society, Washington, DC.

DeLoache, J. S., Pierroutsakos, S. L., Uttal, D. H., Rosengren, K. S., & Gottlieb, A. (1998). Grasping the nature of pictures. *Psychological Science, 9*, 205–210.

Dickie, J. R. (1987). Interrelationships within the mother-father-infant triad. In P. W. Berman & F. A. Pedersen (Eds.), *Men's transitions to parenthood: Longitudinal studies of early family experience* (pp. 113–143). Hillsdale, NJ: Erlbaum.

Diener, E. (2009). Introduction to special issue on the next big questions in psychology. *Perspectives on Psychological Science, 4*, 325. doi: 10.1111/j.1745-6924.2009.0133.x

Diener, E., & Crandall, R. (1978). *Ethics in social and behavioral research.* Chicago: The University of Chicago Press.

Dittmar, H., Halliwell, E., & Ive, S. (2006). Does Barbie make girls want to be thin? The effect of experimental exposure to images of dolls on the body image of 5- to 8-year-old girls. *Developmental Psychology, 42,* 283–292.

Dolan, C. A., Sherwood, A., & Light, K. C. (1992). Cognitive coping strategies and blood pressure responses to real-life stress in healthy young men. *Health Psychology, 11,* 233–240.

Drabick, D. A., & Baugh, D. (2010). A community-based approach to preventing youth violence: What can we learn from the playground? *Progress in Community Health Partnerships: Research, Education, and Action, 4,* 189–196. doi: 10.1353/cpr.2010.0002

Dunlosky, J., Rawson, K. A., Marsh, E. J., Nathan, M. J., & Willingham, D. T. (2013). Improving students' learning with effective learning techniques: Promising directions from cognitive and educational psychology. *Psychological Science in the Public Interest, 14,* 4–58. doi: 10.1177/1529100612453266

Ebneter, D. S., & Latner, J. D. (2013). Stigmatizing attitudes differ across mental health disorders: A comparison of stigma across eating disorders, obesity, and major depressive disorder. *Journal of Nervous & Mental Disease, 201,* 281–285. doi: 10.1097/NMDobo13e318288e23f

Eibl-Eibesfeldt, I. (1975). *Ethology: The biology of behavior.* New York: Holt, Rinehart and Winston.

Eisenberg, D., Downs, M. F., Golberstein, E., & Zivin, K. (2009). Stigma and help seeking for mental health among college students. *Medical Care Research and Review, 66,* 522–541. doi: 10.1177/1077558709335173

Ekman, P. (1994). Strong evidence for universals in facial expressions: A reply to Russell's mistaken critique. *Psychological Bulletin, 115,* 268–287.

Elms, A. C. (2009). Obedience lite. *American Psychologist, 64,* 32–36. doi: 10.1037/a0014473

Endersby, J. W., & Towle, M. J. (1996). Tailgate partisanship: Political and social expression through bumper stickers. *The Social Science Journal, 33,* 307–319.

Entwisle, D. R., & Astone, N. M. (1994). Some practical guidelines for measuring youth's race/ethnicity and socioeconomic status. *Child Development, 65,* 1521–1540.

Epley, N., & Huff, C. (1998). Suspicion, affective response, and educational benefit as a result of deception in psychology research. *Personality and Social Psychology Bulletin, 24,* 759–768.

Epstein, S. (1979). The stability of behavior: On predicting most of the people much of the time. *Journal of Personality and Social Psychology, 37,* 1097–1126.

Ericsson, K. A., & Charness, N. (1994). Expert performance: Its structure and acquisition. *American Psychologist, 49,* 725–747.

Estes, W. K. (1997). On the communication of information by displays of standard errors and confidence intervals. *Psychonomic Bulletin & Review, 4,* 330–341.

Evans, A. D., & Lee, K. (2013, January 7). Emergence of lying in very young children. *Developmental Psychology.* Advance online publication. doi: 10.1037/a0031409

Evans, G. W., Gonnella, C., Marcynyszyn, L. A., Gentile, L., & Salpekar, N. (2005). The role of chaos in poverty and children's socioemotional adjustment. *Psychological Science, 16,* 560–565.

Evans, R., & Donnerstein, E. (1974). Some implications for psychological research of early versus late term participation by college students. *Journal of Research in Personality, 8,* 102–109.

Eysenbach, G., & Till, J. E. (2001, November 10). Ethical issues in qualitative research on Internet communities. *BMJ, 323,* 1103–1105.

Feeney, D. M. (1987). Human rights and animal welfare. *American Psychologist, 42,* 593–599.

Ferdowsian, H. R., Durham, D. L., Kimwele, C., Kranendonk, G., Otali, E., Akugizibwe, T., Mulcahy, J. B., Ajarova, L., & Johnson, C. M. (2011). Signs of mood and anxiety disorders in chimpanzees. *PLoS One, 6*(6), e19855.

Fernandez, K. C., Levinson, C. A., & Rodebaugh, T. L. (2012). Profiling: Predicting social anxiety from Facebook profiles. *Social Psychological and Personality Science*, published online before print, January 19, 2012. doi: 10.1177/1948550611434967

Finch, S., Thomason, N., & Cumming, G. (2002). Past and future American Psychological Association guidelines for statistical practice. *Theory & Psychology, 12*, 825–853.

Fine, M. A., & Kurdek, L. A. (1993). Reflections on determining authorship credit and authorship order on faculty-student collaborations. *American Psychologist, 48*, 1141–1147.

Finkel, E. J., Eastwick, P. W., Karney, B. R., Reis, H. T., & Sprecher, S. (2012). Online dating: A critical analysis from the perspective of psychological science. *Psychological Science in the Public Interest, 13*, 3–66. doi: 10.1177/1529100612436522

Fisher, C. B., & Fryberg, D. (1994). Participant partners: College students weigh the costs and benefits of deceptive research. *American Psychologist, 49*, 417–427.

Fossey, D. (1981). Imperiled giants of the forest. *National Geographic, 159*, 501–523.

Fossey, D. (1983). *Gorillas in the mist*. Boston: Houghton-Mifflin.

Fowler, R. D. (1992). Report of the chief executive officer: A year of building for the future. *American Psychologist, 47*, 876–883.

Fox, J., & Warber, K. M. (2013). Romantic relationship development in the age of Facebook: An exploratory study of emerging adults' perceptions, motives, and behaviors. *Cyberpsychology, Behavior, and Social Networking, 16*, 3–7. doi: 10.1089/cyber.2012.0288

Fraley, R. C. (2004). *How to conduct behavioral research over the Internet*. New York: Guilford Press.

Frick, R. W. (1995). Accepting the null hypothesis. *Memory & Cognition, 23*, 132–138.

Funder, D. C. (2009). Naive and obvious questions. *Perspectives on Psychological Science, 4*, 340–344. doi: 10.1111/j.1745-6924.2009.01135x

Gabrieli, J. D. E., Fleischman, D. A., Keane, M. M., Reminger, S. L., & Morrell, F. (1995). Double dissociation between memory systems underlying explicit and implicit memory in the human brain. *Psychological Science, 6*, 76–82.

Geier, A., Wansink, B., & Rozin, P. (2012). Red potato chips: Segmentation cues can substantially decrease food intake. *Health Psychology, 31*, 398–401. doi: 10.1037/a0027221

Gena, A., Krantz, P. J., McClannahan, L. E., & Poulson, C. L. (1996). Training and generalization of affective behavior displayed by youth with autism. *Journal of Applied Behavioral Analysis, 29*, 291–304.

Gigerenzer, G. (2004). Dread risk, September 11, and fatal traffic accidents. *Psychological Science, 15*, 286–287.

Gilman, R., Connor, N., & Haney, M. (2005). A school-based application of modified habit reversal for Tourette syndrome via a translator: A case study. *Behavior Modification, 29*, 823–838.

Glaser, J., Dixit, J., & Green, D. P. (2002). Studying hate crime with the Internet: What makes racists advocate racial violence? *Journal of Social Issues, 58*, 177–193.

Goodall, J. (1987). A plea for the chimpanzees. *American Scientist, 75*, 574–577.

Gordon, R. T., Schatz, C. B., Myers, L. J., Kosty, M., Gonczy, C., Kroener, J., . . . Zaayer, J. (2008). The use of canines in the detection of human cancers. *Journal of Alternative Complementary Medicine, 14*, 61–67. doi: 10.1089/acm.2006.6408

Gosling, S. (2009). *Snoop: What your stuff says about you*. New York: Basic Books.

Gosling, S. D., & Johnson, J. A. (Eds.). (2010). *Advanced methods for conducting online behavioral research*. Washington, DC: American Psychological Association.

Gosling, S. D., Vazire, S., Srivastava, S., & John, O. P. (2004). Should we trust Web-based studies? A comparative analysis of six preconceptions about Internet questionnaires. *American Psychologist, 59*, 93–104.

Greenwald, A. G., Gonzalez, R., Harris, R. J., & Guthrie, D. (1996). Effect sizes and *p* values: What should be reported and what should be replicated? *Psychophysiology, 33*, 175–183.

Griskevicius, V., Tybur, J. M., & Van den Bergh, B. (2010). Going green to be seen: Status, reputation, and conspicuous conservation. *Journal of Personality and Social Psychology, 98*, 392–404. doi: 10.1037/a0017346

Grissom, R. J., & Kim, J. J. (2005). *Effect sizes for research: A broad practical approach.* Mahwah, NJ: Erlbaum.

Günaydin, G., Zayas, V., Selcuk, E., & Hazan, C. (2012). I like you but I don't know why: Objective facial resemblance to significant others influences snap judgments. *Journal of Experimental Social Psychology, 48*, 350–353. doi: 10.1016/j.jesp.2011.06.001

Hagen, R. L. (1997). In praise of the null hypothesis statistical test. *American Psychologist, 52*, 15–24.

Haggbloom, S. J., Warnick, R., Warnick, J. E., Jones, V. K., Yarbrough, G. L., Russell, T. M., et al. (2002). The 100 most eminent psychologists of the 20th century. *Review of General Psychology, 6*, 139–152.

Halpern, A. R., & Bower, G. H. (1982). Musical expertise and melodic structure in memory for musical notation. *American Journal of Psychology, 95*, 31–50.

Harlow, H. F., & Harlow, M. K. (1966). Learning to love. *American Scientist, 54*, 244–272.

Hart, W., Albarracín, D., Eagly, A. H., Brechan, I., Lindberg, M. J., & Merrill, L. (2009). Feeling validated versus being correct: A meta-analysis of selective exposure to information. *Psychological Bulletin, 135*, 555–588. doi: 10.1037/a0015701

Hartup, W. W. (1974). Aggression in childhood: Development perspectives. *American Psychologist, 29*, 336–341.

Hawkins, D. L., & Pepler, D. J. (2001). Naturalistic observations of peer interventions in bullying. *Social Development, 10*, 512–527. doi: 10.1111/1467-9507.00178

Heatherton, T. F., Mahamedi, F., Striepe, M., Field, A. E., & Keel, P. (1997). A 10-year longitudinal study of body weight, dieting, and eating disorder symptoms. *Journal of Abnormal Psychology, 106*, 117–125.

Heatherton, T. F., Nichols, P., Mahamedi, F., & Keel, P. K. (1995). Body weight, dieting, and eating disorder symptoms among college students 1982 to 1992. *American Journal of Psychiatry, 152*, 1623–1629.

Heatherton, T. F., & Sargent, J. D. (2009). Does watching smoking in movies promote teenage smoking? *Current Directions in Psychological Science, 18*, 63–67. doi: 10.1111/j.1467-8721.2009.01610.x

Hersen, M., & Barlow, D. H. (1976). *Single-case experimental designs: Strategies for studying behavior change.* New York: Pergamon Press.

Hilts, P. J. (1995). Memory's ghost: The nature of memory and the strange tale of Mr. M. New York: Simon & Schuster.

Hippler, H. J., & Schwarz, N. (1987). Response effects in surveys. In H. J. Hippler, N. Schwarz, & S. Sudman (Eds.), *Social information processing and survey methodology* (pp. 102–122). New York: Springer-Verlag.

Hoaglin, D. C., Mosteller, F., & Tukey, J. W. (Eds.). (1983). *Understanding robust and exploratory data analysis.* New York: Wiley.

Hoaglin, D. C., Mosteller, F., & Tukey, J. W. (Eds.). (1991). *Fundamentals of exploratory analysis of variance.* New York: Wiley.

Holden, C. (1987). Animal regulations: So far, so good. *Science, 238*, 880–882.

Holmbeck, G. N. (1997). Toward terminological, conceptual, and statistical clarity in the study of mediators and moderators: Examples from the child-clinical and pediatric psychology literatures. *Journal of Consulting and Clinical Psychology, 65,* 599–610.

Holsti, O. R. (1969). *Content analysis for the social sciences.* Reading, MA: Addison-Wesley.

Horch, J. D., & Hodgins, D. C. (2008). Public stigma of disordered gambling: Social distance, dangerousness, and familiarity. *Journal of Social and Clinical Psychology, 27,* 505–528. doi: 10.1521/jscp.2008.27.5.505

Horton, S. V. (1987). Reduction of disruptive mealtime behavior by facial screening. *Behavior Modification, 11,* 53–64.

Howell, D. C. (2013). *Statistical methods for psychology* (8th ed.). Belmont, CA: Wadsworth.

Hsu, L. M., Chung, J., & Langer, E. J. (2010). The influence of age-related cues on health and longevity. *Perspectives on Psychological Science, 5,* 632–648. doi: 10.1177/1745691610388762

Hunt, M. (1997). *How science takes stock: The story of meta-analysis.* New York: Russell Sage Foundation.

Hunter, J. E. (1997). Needed: A ban on the significance test. *Psychological Science, 8,* 3–7.

Hussain, Z., & Griffiths, M. D. (2009). The attitudes, feelings, and experiences of online gamers: A qualitative analysis. *CyberPsychology & Behavior, 12,* 747–753. doi: 10.1089/cpb.2009.0059

Hyman, I. E., Boss, S. M., Wise, B. M., McKenzie, K. E., & Caggiano, J. M. (2009). Did you see the unicycling clown? Inattentional blindness while walking and talking on a cell phone. *Applied Cognitive Psychology,* published online in Wiley InterScience (www.interscience.wiley.com). doi: 10.1002/acp.1638

Inglehart, R., Foa, R., Peterson, C., & Welzel, C. (2008). Development, freedom, and rising happiness: A global perspective (1981–2007). *Perspectives on Psychological Science, 3,* 264–285. doi: 10.1111/j.1745-6924.2008.00078.x

Johnson, D. (1990). Animal rights and human lives: Time for scientists to right the balance. *Psychological Science, 1,* 213–214.

Jones, D. (2010, June 25). A WEIRD view of human nature skews psychologists' studies. *Science, 328,* 1627.

Jones, N. A., & Bullock, J. J. (2013). Understanding who reported multiple races in the U.S. Decennial Census: Results from census 2000 and the 2010 census. *Family Relations, 62,* 5–16. doi: 10.1111/j.1741-3729.00759.x

Judd, C. M., Smith, E. R., & Kidder, L. H. (1991). *Research methods in social relations* (6th ed.). Fort Worth, TX: Holt, Rinehart and Winston.

Kahneman, D. (2003). A perspective on judgment and choice: Mapping bounded rationality. *American Psychologist, 58,* 697–720.

Kahneman, D. (2011). *Thinking, fast and slow.* New York, NY: Farrar, Straus, and Giroux.

Kahneman, D., & Tversky, A. (1973). On the psychology of prediction. *Psychological Review, 80,* 237–251.

Kaiser, C. R., Vick, S. B., & Major, B. (2006). Prejudice expectations moderate preconscious attention to cues that are threatening to social identity. *Psychological Science, 17,* 332–338.

Kaschak, M. P., & Moore, C. F. (2000). On the documentation of statistical analyses in the "Clicky-Box" era. *American Psychologist, 55,* 1511–1512.

Kassin, S. M., Goldstein, C. C., & Savitsky, K. (2003). Behavioral confirmation in the interrogation room: On the dangers of presuming guilt. *Law and Human Behavior, 27,* 187–203.

Kassin, S. M., & Kiechel, K. L. (1996). The social psychology of false confessions: Compliance, internalization, and confabulation. *Psychological Science, 7,* 125–128.

Kazdin, A. E. (1978). Methodological and interpretive problems of single-case experimental designs. *Journal of Consulting and Clinical Psychology, 46,* 629–642.

Kazdin, A. E. (1980). *Behavior modification in applied settings* (rev. ed.). Homewood, IL: Dorsey Press.

Kazdin, A. E. (1982). Single-case experimental designs. In P. C. Kendall & J. N. Butcher (Eds.), *Handbook of research methods in clinical psychology* (pp. 416–490). New York: Wiley.

Kazdin, A. (2011). *Single-case research designs: Methods for clinical and applied settings.* New York: Oxford University Press.

Keel, P. K., Baxter, M. G., Heatherton, T. F., & Joiner, T. E., Jr. (2007). A 20-year longitudinal study of body weight, dieting, and eating disorder symptoms. *Journal of Abnormal Psychology, 116,* 422–432. doi: 10.1037/0021-843X.116.2.422

Kelman, H. C. (1967). Human use of human subjects: The problem of deception in social psychological experiments. *Psychological Bulletin, 67,* 1–11.

Kelman, H. C. (1972). The rights of the subject in social research: An analysis in terms of relative power and legitimacy. *American Psychologist, 27,* 989–1016.

Kenny, D. A. (1979). *Correlation and causality.* New York: Wiley.

Keppel, G. (1991). *Design and analysis: A researcher's handbook* (3rd ed.). Englewood Cliffs, NJ: Prentice-Hall.

Khuder, S. A., Milz, S., Jordan, T., Price, J., Silvestri, K., & Butler, P. (2007). The impact of a smoking ban on hospital admissions for coronary heart disease. *Preventive Medicine, 45,* 3–8. doi: 10.1016/j/ypmed.2007.03.011

Kidd, S. A. (2002). The role of qualitative research in psychological journals. *Psychological Methods, 7,* 126–138.

Kidd, S. A., & Kral, M. J. (2002). Suicide and prostitution among street youth: A qualitative analysis. *Adolescence, 37,* 411–430.

Kimble, G. A. (1989). Psychology from the standpoint of a generalist. *American Psychologist, 44,* 491–499.

Kimmel, A. J. (1996). *Ethical issues in behavioral research: A survey.* Cambridge, MA: Blackwell.

Kimmel, A. J. (1998). In defense of deception. *American Psychologist, 53,* 803–805.

Kirk, R. E. (1996). Practical significance: A concept whose time has come. *Educational and Psychological Measurement, 56,* 746–759.

Kirkham, G. L. (1975). Doc cop. *Human Behavior, 4,* 16–23.

Kirsch, I. (1978). Teaching clients to be their own therapists: A case-study illustration. *Psychotherapy: Theory, Research and Practice, 15,* 302–305.

Kirsch, I., & Sapirstein, G. (1998). Listening to Prozac but hearing placebo: A meta-analysis of antidepressant medication. *Prevention & Treatment, 1*(2). doi: 10.1037/1522-3736.1.1.12a, prevention/volume1/pre0010002a.html.

Klein, E. (2013a, April 28). If this were a pill, you'd do anything to get it. *The Washington Post,* G1, G6–G7.

Klein, E. (2013b, May 5). Oregon's Medicaid lottery should become a case study in smarter spending. *The Washington Post,* G1, G3.

Klinesmith, J., Kasser, T., & McAndrew, F. T. (2006). Guns, testosterone, and aggression: An experimental test of a meditational hypothesis. *Psychological Science, 17,* 568–571.

Knafo, A., Schwartz, S. H., & Levine, R. V. (2009). Helping strangers is lower in embedded cultures. *Journal of Cross-Cultural Psychology, 40,* 875–879. doi: 10.1177/0022022109339211

Kramer, A. D. I. (2010). An unobtrusive behavioral model of "Gross National Happiness." Proceedings of the 28th Annual Conference on Human Factors in Computing Systems, April 10–15, Atlanta, Georgia, USA, 287–290.

Krantz, J. H., & Dalal, R. (2000). Validity of Web-based psychological research. In M. H. Birnbaum (Ed.), *Psychological experiments on the Internet* (pp. 35–60). San Diego, CA: Academic Press.

Kratochwill, T. R., & Levin, J. R. (Eds.). (1992). *Single-case research designs and analysis: New directions for psychology and education*. Mahwah, NJ: Erlbaum.

Kratochwill, T. R., & Martens, B. K. (1994). Applied behavior analysis and school psychology. *Journal of Applied Behavior Analysis, 27*, 3–5.

Kraut, R., Olson, J., Banaji, M. R., Bruckman, A., Cohen, J., & Couper, M. (2004). Psychological research online: Report of Board of Scientific Affairs' Advisory Group on the conduct of research on the Internet. *American Psychologist, 59*, 105–117.

Krueger, J. (2001). Null hypothesis significance testing: On the survival of a flawed method. *American Psychologist, 56*, 16–26.

Kruger, J., Wirtz, D., & Miller, D. (2005). Counterfactual thinking and the first instinct fallacy. *Journal of Personality and Social Psychology, 88*, 725–735. doi: 10.1037/0022-3514.88.5.725

Kubany, E. S. (1997). Application of cognitive therapy for trauma-related guilt (CT-TRG) with a Vietnam veteran troubled by multiple sources of guilt. *Cognitive and Behavioral Practice, 4*, 213–244.

Lambert, N. M., Clark, M. S., Durtschi, J., Fincham, F. D., & Graham, S. M. (2010). Benefits of expressing gratitude: Expressing gratitudes to a partner changes one's view of the relationship. *Psychological Science, 21*, 574–580. doi: 10.1177/09567976103644003

Lang, R., Regester, A., Mulloy, A., Rispoli, M, & Botout, A. (2011). Behavioral intervention to treat selective mutism across multiple social situations and community settings. *Journal of Applied Behavior Analysis, 44*, 623–628. doi: 10.1901/jaba.2011.44-623

Langer, E. J. (1989). *Mindfulness*. Reading, MA: Addison-Wesley.

Langer, E. J. (1997). *The power of mindful learning*. Reading, MA: Addison-Wesley.

Langer, E. J., & Piper, A. I. (1987). The prevention of mindlessness. *Journal of Personality and Social Psychology, 53*, 280–287.

Langer, E. J., & Rodin, J. (1976). The effects of choice and enhanced personal responsibility for the aged: A field experiment in an institutional setting. *Journal of Personality and Social Psychology, 34*, 191–198.

Larson, R. (1989). Beeping children and adolescents: A method for studying time use and daily experience. *Journal of Youth and Adolescence, 18*, 511–530.

Larson, R. W. (2000). Toward a psychology of positive youth development. *American Psychologist, 55*, 170–183. doi: 10.1037/0003-066x.55.1.170

Larson, R. W. (2011). Positive development in a disorderly world. *Journal of Research on Adolescence, 21*, 317–334. doi: 10.1111/j.1532-7795.2010.00707.x

Larson, R. W., Richards, M. H., Moneta, G., Holmbeck, G., & Duckett, E. (1996). Changes in adolescents' daily interactions with their families from ages 10 to 18: Disengagement and transformation. *Developmental Psychology, 32*, 744–754.

Latané, B., & Darley, J. M. (1970). *The unresponsive bystander: Why doesn't he help?* New York: Appleton-Century-Crofts.

LeBlanc, P. (2001, September). "And mice." (Or tips for dealing with the animal subjects review board). *APS Observer, 14*, 21–22.

Lee, S. W. S., Schwarz, N., Taubman, D., & Hou, M. (2010). Sneezing in times of a flu pandemic: Public sneezing increases perception of unrelated risks and shifts preferences for federal spending. *Psychological Science, 21*, 375–377. doi: 10.1177/ 0956797609359876

Levine, R. V., Norenzayan, A., & Philbrick, K. (2001). Cross-cultural differences in helping strangers. *Journal of Cross-Cultural Psychology, 32*, 543–560. doi: 10.1177/0022022101032005002

Levitt, S. D., & Dubner, S. J. (2005). *Freakonomics: A rogue economist explores the hidden side of everything*. New York: HarperCollins.

Levitt, S. D., & Dubner, S. J. (2009). *SuperFreakonomics: Global cooling, patriotic prostitutes, and why suicide bombers should buy life insurance*. New York: HarperCollins.

Li, M., Vietri, J., Galvani, A. P., & Chapman, G. B. (2010). How do people value life? *Psychological Science, 21*, 163–167. doi: 10.1177/0956797609357707

Lilienfeld, S. O. (2010). Can psychology become a science? *Personality and Individual Differences, 49*, 281–288. doi: 10.1016/j.paid.2010.01.024

Lilienfeld, S. O. (2012). Public skepticism of psychology: Why many people perceive the study of human behavior as unscientific. *American Psychologist, 67*, 111–129. doi: 10.1037/a0023963

Locke, T. P., Johnson, G. M., Kirigin-Ramp, K., Atwater, J. D., & Gerrard, M. (1986). An evaluation of a juvenile education program in a state penitentiary. *Evaluation Review, 10*, 281–298.

Loftus, E. F. (2003, August). Loftus: The need to defend scientific freedom. *APS Observer, 16*, 1, 32.

Loftus, G. R. (1991). On the tyranny of hypothesis testing in the social sciences. *Contemporary Psychology, 36*, 102–105.

Loftus, G. R. (1996). Psychology will be a much better science when we change the way we analyze data. *Current Directions in Psychological Science, 5*, 161–171.

Loftus, G. R., & Masson, M. E. J. (1994). Using confidence intervals in within-subject design. *Psychonomic Bulletin & Review, 1*, 476–490.

Lopresti-Goodman, S. M., Kameka, M., & Dube, A. (2013). Stereotypical behaviors in chimpanzees rescued from the African bushmeat and pet trade. *Behavioral Sciences, 3*, 1–20. doi: 10.3390/bs3010001

Lucas, R. E., Diener, E., & Suh, E. (1996). Discriminant validity of well-being measures. *Journal of Personality and Social Psychology, 71*, 616–628.

MacCoun, R. (2002, December). Why a psychologist won the Nobel Prize in economics. *APS Observer, 15*, 1, 8.

Marx, M. H. (1963). The general nature of theory construction. In M. H. Marx (Ed.), *Theories in contemporary psychology* (pp. 4–46). New York: Macmillan.

Matsumoto, D., & Willingham, B. (2006). The thrill of victory and the agony of defeat: Spontaneous expressions of medal winners of the 2004 Athens Olympic Games. *Journal of Personality and Social Psychology, 91*, 568–581.

Mayer, R. E., Stull, A., DeLeeuw, K., Almeroth, K., Bimber, B., Chun, D., . . . Zhang, H. (2009). Clickers in college classrooms: Fostering learning with questioning methods in large lecture classes. *Contemporary Educational Psychology, 34*, 51–57. doi: 10.1016/j.cedpsych.2008.04.002

Mazur, E. (2010). Collecting data from social networking Web sites and blogs. In S. D. Gosling & J. A. Johnson (Eds.), *Advanced methods for conducting online behavioral research* (pp. 77–90). Washington, DC: American Psychological Association.

McCallum, D. M. (2001, May/June). "Of men . . ." (Or how to obtain approval from the human subjects review board). *APS Observer, 14*, 28–29, 35.

McCarthy, A., Lee, K., Itakura, S., & Muir, D. W. (2006). Cultural display rules drive eye gaze during thinking. *Journal of Cross-Cultural Psychology, 37*, 717–722.

McCarthy, D. E., Piasecki, T., M., Fiore, M. C., & Baker, T. B. (2006). Life before and after quitting smoking: An electronic diary study. *Journal of Abnormal Psychology, 115*, 454–466.

McCulloch, M., Jezierski, T., Broffman, M., Hubbard A., Turner, K., & Janecki, T. (2006). Diagnostic accuracy of canine scent detection in early- and late-stage lung and breast cancers. *Integrative Cancer Therapies, 5*, 30–39. doi: 10.1177/1534735405285096

McGuire, W. J. (1997). Creative hypothesis generating in psychology: Some useful heuristics. *Annual Review of Psychology, 48,* 1–30.

Meehl, P. E. (1967). Theory-testing in psychology and physics: A methodological. *Philosophy of Science, 34,* 103–115.

Meehl, P. E. (1978). Theoretical risks and tabular asterisks: Sir Karl, Sir Ronald, and the slow progress of soft psychology. *Journal of Consulting and Clinical Psychology, 46,* 806–834.

Meehl, P. E. (1990a). Appraising and amending theories: The strategy of Lakatosian defense and two principles that warrant it. *Psychological Inquiry, 1,* 108–141.

Meehl, P. E. (1990b). Why summaries of research on psychological theories are often uninterpretable. *Psychological Reports, 66,* 195–244 (Monograph Supplement 1-V66).

Mehl, M. R., & Holleran, S. E. (2007). An empirical analysis of the obtrusiveness of and participants' compliance with the electronically activated recorder (EAR). *European Journal of Psychological Assessment, 23,* 248–257. doi: 10.1027/1015-5759.23.4.248

Mehl, M. R., & Pennebaker, J. W. (2003). The sounds of life: A psychometric analysis of students' daily social environments and natural conversations. *Journal of Personality and Social Psychology, 84,* 857–870. doi:10.1037/0022-3514.84.4.857

Mehl, M. R., Robbins, M. L., & Deters, F. G. (2012). Naturalistic observation of health-relevant social processes: The electronically activated recorder methodology in psychosomatics. *Psychosomatic Medicine, 74,* 410–417. doi: 10.1097/PSY.0b013e3182545470

Mehl, M. R., Vazire, S., Holleran, S. E., & Clark, C. S. (2010). Eavesdropping on happiness: Well-being is related to having less small talk and more substantive conversations. *Psychological Science, 21,* 539–541. doi: 10.1177/0956797610362675

Mehl, M. R., Vazire, S., Ramírez-Esparza, N., Slatcher, R. B., & Pennebaker, J. W. (2007, July 6). Are women really more talkative than men? *Science, 317,* 82. doi: 10.1126/science.1139940

Michielutte, R., Shelton, B., Paskett, E. D., Tatum, C. M., & Velez, R. (2000). Use of an interrupted time-series design to evaluate a cancer screening program. *Health Education Research, 15,* 615–623.

Miles, M. B., & Huberman, A. M., & Saldaña, J. (2013). *Qualitative data analysis: A methods sourcebook* (3rd ed.). Thousand Oaks, CA: Sage.

Milgram, S. (1963). Behavioral study of obedience. *Journal of Abnormal and Social Psychology, 67,* 371–378. doi: 10.1037/h0040525

Milgram, S. (1965). Some conditions of obedience and disobedience to authority. *Human Relations, 18,* 57–76. doi: 10.1177/001872676501800105

Milgram, S. (1974). *Obedience to authority.* New York: Harper & Row.

Milgram, S. (1977, October). Subject reaction: The neglected factor in the ethics of experimentation. *Hastings Center Report.*

Miller, G. A., & Wakefield, P. C. (1993). On Anglin's analysis of vocabulary growth. *Monographs of the Society for Research in Child Development, 58* (10, Serial No. 238).

Miller, N. E. (1985). The value of behavioral research on animals. *American Psychologist, 40,* 423–440.

Mooallem, J. (2010, April 4). The love that dare not squawk its name: Inside the science of same-sex animal pairings. *The New York Times Magazine,* pp. 26–35, 44, 46.

Mook, D. G. (1983). In defense of external invalidity. *American Psychologist, 38,* 379–387.

Mosteller, F., & Hoaglin, D. C. (1991). Preliminary examination of data. In D. C. Hoaglin, F. Mosteller, & J. W. Tukey (Eds.), *Fundamentals of exploratory analysis of variance* (pp. 40–49). New York: Wiley.

Mulaik, S. A., Raju, N. S., & Harshman, R. A. (1997). There is a time and place for significance testing. In L. L. Harlow, S. A. Mulaik, & J. H. Steiger (Eds.), *What if there were no significance tests?* (pp. 65–115). Mahwah, NJ: Erlbaum.

Myers, D. G., & Diener, E. (1995). Who is happy? *Psychological Science, 6,* 10–19.

National Academy of Sciences. (2011). *Guide for the Care and Use of Laboratory Animals* (8th ed.). Report of the Committee for the Upgrade of the *Guide*, Janet C. Garber (Chair). Washington, DC: National Academies Press.

Neisser, U. (1967). *Cognitive psychology*. New York: Appleton-Century-Crofts.

Neisser, U., & Harsch, N. (1992). Phantom flashbulbs: False recollections of hearing the news about *Challenger*. In E. Winograd & U. Neisser (Eds.), *Affect and accuracy in recall: Studies of "flashbulb memories"* (pp. 9–31). New York: Cambridge University Press.

Newburger, E. C. (2001, September, U.S. Census Bureau). *Home computers and Internet use in the United States: August 2000*. Retrieved June 1, 2004, from http://www.census.gov/prod/2001pubs/p23-207.pdf

Newhagen, J. E., & Ancell, M. (1995). The expression of emotion and social status in the language of bumper stickers. *Journal of Language and Social Psychology, 14*, 312–323.

Newman, E. (2008, May). Assessing trauma and its effects without distress: A guide to working with IRBs. *APS Observer, 21*(5). Retrieved from www.psychologicalscience.org/index.php/publications/observer/2008/may-08/assessing-trauma-and-its-effects-without-distress-a-guide-to-working-with-irbs.html

Nosek, B. A., Banaji, M. R., & Greenwald, A. G. (2002). E-Research: Ethics, security, design, and control in psychological research on the Internet. *Journal of Social Issues, 58*, 161–176.

Novak, M. A. (1991, July). "Psychologists care deeply" about animals. *Monitor on Psychology*, 4.

Ondersma, S. J., Chaffin, M., Berliner, L., Cordon, I., Goodman, G. S., & Barnett, D. (2001). Sex with children is abuse: Comment on Rind, Tromovitch, and Bauserman (1998). *Psychological Bulletin, 127*, 707–714.

Orne, M. T. (1962). On the social psychology of the psychological experiment: With particular reference to demand characteristics and their implications. *American Psychologist, 17*, 776–783.

Ortmann, A., & Hertwig, R. (1997). Is deception necessary? *American Psychologist, 52*, 746–747.

Ostrov, J. M., & Keating, C. F. (2004). Gender differences in preschool aggression during free play and structured interactions: An observational study. *Social Development, 13*, 255–277. doi: 10.1111/j.1467-9507.2004.000266.x

Park, C. L., Armeli, S., & Tennen, H. (2004). Appraisal-coping goodness of fit: A daily Internet study. *Personality and Social Psychology Bulletin, 30*, 558–569.

Parker, R. I., & Brossart, D. F. (2003). Evaluating single-case research data: A comparison of seven statistical methods. *Behavior Therapy, 34*, 189–211. doi: 10.1016/S0005-7893(03)80013-8

Parry, H. J., & Crossley, H. M. (1950). Validity of responses to survey questions. *Public Opinion Quarterly, 14*, 61–80.

Parsons, H. M. (1974). What happened at Hawthorne? *Science, 183*, 922–932.

Parsonson, B. S., & Baer, D. M. (1992). The visual analysis of data, and current research into the stimuli controlling it. In T. R. Kratochwill & J. R. Levin (Eds.), *Single-case research design and analysis* (pp. 15–40). Hillsdale, NJ: Erlbaum.

Pashler, H., & Wagenmakers, E. J. (2012). Editors' introduction to the special section on replicability in psychological science: A crisis of confidence? *Perspectives on Psychological Science, 7*, 528–530. doi: 10.1177/1745691612465253

Pease, A., & Pease, B. (2006). *The definitive book of body language*. New York: Bantam Dell.

Pennebaker, J. W. (1989). Confession, inhibition, and disease. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 22, pp. 211–244). New York: Academic Press.

Pennebaker, J. W., & Francis, M. E. (1996). Cognitive, emotional, and language processes in disclosure. *Cognition and Emotion, 10,* 601–626.

Peterson, C., & Seligman, M. E. P. (2003). Character strengths before and after September 11. *Psychological Science, 14,* 381–384. doi: 10.1111/1467-9280.24482

Piaget, J. (1965). *The child's conception of number.* New York: Norton.

Pickren, W. E. (2003). An elusive honor: Psychology, behavior, and the Nobel Prize. *American Psychologist, 58,* 721–722.

Pingitore, R., Dugoni, B. L., Tindale, R. S., & Spring, B. (1994). Bias against overweight job applicants in a simulated employment interview. *Journal of Applied Psychology, 79,* 909–917.

Pollick, A. (2007, November). IRBs: Navigating the maze. *APS Observer, 20*(10). Retrieved from www.psychologicalscience.org/index.php/publications/observer/2008/november-07/irbs-navigating-the-maze.html

Popper, K. R. (1959). *The logic of scientific discovery.* New York: Basic Books.

Popper, K. R. (1976). *Unended quest.* Glasgow: Fontana/Collins.

Posavac, E. J. (2002). Using *p* values to estimate the probability of a statistically significant replication. *Understanding Statistics, 1,* 101–112.

Posavac, E. J. (2011). *Program evaluation* (8th ed.). Englewood Cliffs, NJ: Prentice-Hall.

Poulton, E. C. (1973). Unwanted range effects from using within-subject experimental designs. *Psychological Bulletin, 80,* 113–121.

Poulton, E. C. (1975). Range effects in experiments on people. *American Journal of Psychology, 88,* 3–32.

Poulton, E. C. (1982). Influential companions. Effects of one strategy on another in the within-subjects designs of cognitive psychology. *Psychological Bulletin, 91,* 673–690.

Poulton, E. C., & Freeman, P. R. (1966). Unwanted asymmetrical transfer effects with balanced experimental designs. *Psychological Bulletin, 66,* 1–8.

Powell, K. L., Roberts, G., & Nettle, D. (2012). Eye images increase charitable donations: Evidence from an opportunistic field experiment in a supermarket. *Ethology, 118,* 1–6. doi: 10.1111/eth.12011

Price, M. (2010, July/August). Sins against science: Data fabrication and other forms of scientific misconduct may be more prevalent than you think. *Monitor on Psychology,* 41(7). Retrieved from http://www.apa.org/monitor/2010/07-08/misconduct.aspx

Pryor, J. H., Eagan, K., Blake, L. P., Berdan, J., & Case, M. H. (2012). *The American freshman: National norms fall 2012.* Los Angeles: Higher Education Research Institute, UCLA.

Pryor, J. H., Hurtado, S., DeAngelo, L., Patuki Blake, L., & Tran, S. (2009). *The American freshman: National norms fall 2009.* Los Angeles: Higher Education Research Institute, UCLA.

Ramírez-Esparza, N., Mehl, M. R., Álvarez-Bermúdez, J., & Pennebaker, J. W. (2009). Are Mexicans more or less sociable than Americans? Insights from a naturalistic observation study. *Journal of Research in Personality, 43,* 1–7. doi: 10.1016/j.jrp.2008.09.002

Rasinski, K. A., Willis, G. B., Baldwin, A. K., Yeh, W., & Lee, L. (1999). Methods of data collection, perceptions of risks and losses, and motivation to give truthful answers to sensitive survey questions. *Applied Cognitive Psychology, 13,* 465–484.

Rauscher, F. H., Shaw, G. L., & Ky, K. N. (1993). Music and spatial task performance. *Nature, 365,* 611.

Richardson, D. R., Pegalis, L., & Britton, B. (1992). A technique for enhancing the value of research participation. *Contemporary Social Psychology, 16,* 11–13.

Richardson, J. & Parnell, P. (2005). *And Tango makes three.* Simon Schuster.

Riley, D. A. (1962). Memory for form. In L. Postman (Ed.), *Psychology in the making* (pp. 402–465). New York: Knopf.

Rimm, D. C., & Masters, J. C. (1979). *Behavior therapy: Techniques and empirical findings* (2nd ed.). New York: Academic Press.

Rind, B., & Tromovitch, P. (2007). National samples, sexual abuse in childhood, and adjustment in adulthood: A commentary on Najman, Dunne, Purdie, Boyle, and Coxeter (2005). *Archives of Sexual Behavior, 36,* 101–106, doi: 10.1007/s10508-006-9058-y

Rind, B., Tromovitch, P., & Bauserman, R. (1998). A meta-analytic examination of assumed properties of child sexual abuse using college samples. *Psychological Bulletin, 124,* 22–53.

Rind, B., Tromovitch, P., & Bauserman, R. (2001). The validity and appropriateness of methods, analyses, and conclusions in Rind et al. (1998): A rebuttal of victimological critique from Ondersma et al. (2001) and Dallam et al. (2001). *Psychological Bulletin, 127,* 734–758.

Robins, R. W., Gosling, S. D., & Craik, K. H. (1999). An empirical analysis of trends in psychology. *American Psychologist, 54,* 117–128.

Roethlisberger, F. J. (1977). *The elusive phenomena: An autobiographical account of my work in the field of organized behavior at the Harvard Business School.* Cambridge, MA: Division of Research, Graduate School of Business Administration (distributed by Harvard University Press).

Rogers, A. (1999). *Barbie culture.* Thousand Oaks, CA: Sage.

Rosenhan, D. L. (1973). On being sane in insane places. *Science, 179,* 250–258.

Rosenthal, R. (1963). On the social psychology of the psychological experiment: The experimenter's hypothesis as unintended determinant of experimental results. *American Scientist, 51,* 268–283.

Rosenthal, R. (1966). *Experimenter effects in behavioral research.* New York: Appleton-Century-Crofts.

Rosenthal, R. (1976). *Experimenter effects in behavioral research.* (Enlarged ed.). New York: Irvington.

Rosenthal, R. (1990). How are we doing in soft psychology? *American Psychologist, 45,* 775–777.

Rosenthal, R. (1991). *Meta-analytic procedures for social research* (Rev. ed.). Newbury Park, CA: Sage.

Rosenthal, R. (1994a). Interpersonal expectancy effects: A 30-year perspective. *Current Directions in Psychological Science, 3,* 176–179.

Rosenthal, R. (1994b). Science and ethics in conducting, analyzing, and reporting psychological research. *Psychological Science, 5,* 127–134.

Rosenthal, R., & Rosnow, R. L. (1991). *Essentials of behavioral research: Methods and data analysis* (2nd ed.). New York: McGraw-Hill.

Rozin, P. (2009). What kind of empirical research should we publish, fund, and reward? A different perspective. *Perspectives on Psychological Science, 4,* 435–439. doi: 10.1111/j.1745-6924.2009.01151.x

Rozin, P., Kabnick, K., Pete, E., Fischler, C., & Shields, C. (2003). The ecology of eating: Smaller portion sizes in France than in the United States help explain the French paradox. *Psychological Science, 14,* 450–454.

Rule, N. O., Krendl, A. C., Ivcevic, Z., & Ambadi, N. (2013). Accuracy and consensus in judgments of trustworthiness from faces: Behavioral and neural correlates. *Journal of Personality and Social Psychology, 104,* 409–426. doi: 10.1037/a0031050

Sacks, O. (2012). *Hallucinations.* New York: Knopf.

Sacks, O. (1985). *The man who mistook his wife for a hat and other clinical tales.* New York: Harper & Row.

Sacks, O. (1995). *An anthropologist on Mars.* New York: Knopf.

Sacks, O. (2007). *Musicophilia: Tales of music and the brain.* New york: A. A. Knopf.

Saldaña, D. (2013). *The coding manual for qualitative researchers* (3rd ed.).Thousand Oaks, CA: Sage.

Salomon, G. (1987). Basic and applied research in psychology: Reciprocity between two worlds. *International Journal of Psychology, 22,* 441–446.

Sax, L. J., Astin, A. W., Lindholm, J. A., Korn, W. S., Saenz, V. B., & Mahoney, K. M. (2003). *The American freshman: National norms for fall 2003.* Los Angeles: Higher Education Research Institute, UCLA.

Schacter, D. L. (1996). *Searching for memory.* New York: Basic Books.

Schaus, J. F., Sole, M. L., McCoy, T. P., Mullett, N., & O'Brien, M. C. (2010). Alcohol screening and brief intervention in a college student health center: A randomized controlled trial. *Journal of Studies on Alcohol and Drugs, Supplement,* (Supplement no. 16), 131–141.

Schellenberg, E. G. (2005). Music and cognitive abilities. *Current Directions in Psychological Science, 14,* 317–320. doi: 10.1111/j.0963-7214.2005.00389.x

Scherr, K. C., & Madon, S. (2013). "Go ahead and sign": An experimental examination of Miranda waivers and comprehension. *Law and Human Behavior, 37,* 208–218. doi: 10.1037/lhb0000026

Schmidt, F. L. (1996). Statistical significance testing and cumulative knowledge in psychology: Implications for training of researchers. *Psychological Methods, 1,* 115–129.

Schmidt, F. L., & Hunter, J. E. (1997). Eight common but false objections to the discontinuation of significance testing in the analysis of research data. In L. L. Harlow, S. A. Mulaik, & J. H. Steiger (Eds.), *What if there were no significance tests?* (pp. 37–64). Mahwah, NJ: Erlbaum.

Schmidt, W. C. (1997). World-Wide-Web survey research: Benefits, potential problems, and solutions. *Behavior Research Methods, Instruments, & Computers, 29,* 274–279.

Schoeneman, T. J., & Rubanowitz, D. E. (1985). Attributions in the advice columns: Actors and observers, causes and reasons. *Personality and Social Psychology Bulletin, 11,* 315–325.

Schwartz, P. (2010, January/February). Love, American style. *The AARP Magazine.*

Scoville, W. B., & Milner, B. (1957). Loss of recent memory after bilateral hippocampal lesions. *Journal of Neurology, Neurosurgery, and Psychiatry, 20,* 11–19.

Seligman, M. E. P., Steen, T. A., Park, N., & Peterson, C. (2005). Positive psychology progress: Empirical validation of interventions. *American Psychologist, 60,* 410–421.

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference.* Boston: Houghton Mifflin.

Shapiro, K. J. (1998). *Animal models of human psychology: Critique of science, ethics, and policy.* Seattle, WA: Hogrefe & Huber.

Shiffman, S., & Paty, J. (2006). Smoking patterns and dependence: Contrasting chippers and heavy smokers. *Journal of Abnormal Psychology, 115,* 509–523.

Silverman, D. (2011). *Qualitative research* (3rd ed.). Thousand Oaks, CA: Sage.

Simon, H. A. (1992). What is an "explanation" of behavior? *Psychological Science, 3,* 150–161.

Singer, P. (1990). The significance of animal suffering. *Behavioral and Brain Sciences, 13,* 9–12.

Singer, P. (2011). *Practical ethics* (3rd ed.). New York: Cambridge University Press.

Skinner, B. F. (1966). Operant behavior. In W. K. Honig (Ed.), *Operant behavior: Areas of research and application* (pp. 12–32). New York: Appleton-Century-Crofts.

Skitka, L. J., Bauman, C. W., & Mullen, E. (2004). Political tolerance and coming to psychological closure following the September 11, 2001, terrorist attacks: An integrative approach. *Personality and Social Psychology Bulletin, 30,* 743–756. doi: 10.1177/0146167204263968

Skitka, L. J., & Sargis, E. G. (2005). Social psychological research and the Internet: The promise and the perils of a new methodological frontier. In Y. Amichai-Hamburger (Ed.), *The social net: The social psychology of the Internet*. New York: Oxford University Press.

Slatcher, R. B., & Trentacosta, C. J. (2012). A naturalistic study of the links between parental depressive symptoms and preschoolers' behaviors in everyday life. *Journal of Family Therapy, 25*, 444–448. doi: 10.1037/a0023728

Smith, J. A., Harré, R., & Van Langenhove, L. (1995). Idiography and the case study. In J. A. Smith, R. Harré, & L. Van Langenhove (Eds.), *Rethinking psychology* (pp. 59–69). Thousand Oaks, CA: Sage.

Smith, T. W. (1981). Qualifications to generalized absolutes: "Approval of hitting" questions on the GSS. *Public Opinion Quarterly, 45*, 224–230.

Snibbe, A. C., & Markus, H. R. (2005). You can't always get what you want: Educational attainment, agency, and choice. *Journal of Personality and Social Psychology, 88*, 703–720. doi: 10.1037/0022-3514.88.4.703

Sokal, M. M. (1992). Origins and early years of the American Psychological Association, 1890–1906. *American Psychologist, 47*, 111–122.

Solomon, Z., Gelkopf, M., & Bleich, A. (2005). Is terror gender-blind? Gender differences in reaction to terror events. *Social Psychiatry and Psychiatric Epidemiology, 40*, 947–954. doi:10.1007/s00127-005-0973-3

Spitz, R. A. (1965). *The first year of life*. New York: International Universities Press.

Spitzer, R. L. (1976). More on pseudoscience in science and the case for psychiatric diagnosis. *Archives of General Psychiatry, 33*, 459–470.

Steele, K. M., Bass, K. E., & Crook, M. D. (1999). The mystery of the Mozart effect: Failure to replicate. *Psychological Science, 10*, 366–369. doi: 10.1111/1467-9289.00169

Stephens, N. M., Hamedani, M. G., Markus, H. R., Bergsieker, H. B., & Eloul, L. (2009). Why did they "choose" to stay? Perspectives of Hurricane Katrina observers and survivors. *Psychological Science, 20*, 878–886. doi: 10.1111/j.1467-9280.2009.02386.x

Sternberg, R. J. (1997, September). What do students still most need to learn about research in psychology? *APS Observer*, 14, 19.

Subrahmanyam, K., Greenfield, P. M., & Tynes, B. (2004). Constructing sexuality and identity in an online teen chat room. *Journal of Applied Developmental Psychology, 25*, 651–666. doi: 10.1016/j.appdev.2004.09.007

Sue, S. (1999). Science, ethnicity and bias. *American Psychologist, 54*, 1070–1077.

Surwit, R. S., & Williams, P. G. (1996). Animal models provide insight into psychosomatic factors in diabetes. *Psychosomatic Medicine, 58*, 582–589.

Susskind, J. E. (2003). Children's perception of gender-based illusory correlations: Enhancing preexisting relationships between gender and behavior. *Sex Roles, 48*, 483–494.

Swami, V. (2012). Written on the body? Individual differences between British adults who do and do not obtain a first tattoo. *Scandinavian Journal of Psychology, 53*, 407–412. doi: 10.1111/j.1467-9450.2012.00960.x

Talarico, J. M., & Moore, K. M. (2012). Memories of "The Rivalry": Differences in how fans of winning and losing teams remember the same game. *Applied Cognitive Psychology, 26*, 746–756. doi: 10.1002/acp.2855

Taylor, K. M., & Shepperd, J. A. (1996). Probing suspicion among participants in deception research. *American Psychologist, 51*, 886.

Thioux, M., Stark, D. E., Klaiman, C., & Schultz, R. T. (2006). The day of the week when you were born in 700 ms: Calendar computation in an autistic savant. *Journal of Experimental Psychology: Human Perception and Performance, 32*, 1155–1168.

Thomas, L. (1992). *The fragile species*. New York: Charles Scribner's Sons.

Thompson, W. F., Schellenberg, E. G., & Husain, G. (2001). Arousal, mood, and the Mozart effect. *Psychological Science, 12*, 248–251. doi: 10.1111/1467-9280.00345

Tukey, J. W. (1977). *Exploratory data analysis.* Reading, MA: Addison-Wesley.
Tuten, T. L. (2010). Conducting online surveys. In S. D. Gosling & J. A. Johnson (Eds.), *Advanced methods for conducting online behavioral research* (pp. 179–192). Washington, D.C.: American Psychological Association.
Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science, 185,* 1124–1131.
Ulrich, R. E. (1991). Animal rights, animal wrongs and the question of balance. *Psychological Science, 2,* 197–201.
Ulrich, R. E. (1992). Animal research: A reflective analysis. *Psychological Science, 3,* 384–386.
Underwood, B. J., & Shaughnessy, J. J. (1975). *Experimentation in psychology.* New York: Wiley: Robert E. Krieger, 1983.
Utz, S. (2010). Using automated "field notes" to observe the behavior of online subjects. In S. D. Gosling & J. A. Johnson (Eds.), *Advanced methods for conducting online behavioral research* (pp. 91–108). Washington, D.C.: American Psychological Association.
Valentino, K., Cicchetti, D., Toth, S. L., & Rogosch, F. A. (2006). Mother-child play and emerging social behaviors among infants from maltreating families. *Developmental Psychology, 42,* 474–485.
van Baaren, R. B., Holland, R. W., Kawakami, K., & van Knippenberg, A. (2004). Mimicry and prosocial behavior. *Psychological Science, 15,* 71–74.
VanderStoep, S. W., & Shaughnessy, J. J. (1997). Taking a course in research methods improves reasoning about real-life events. *Teaching of Psychology, 24,* 122–124.
Walther, J. B., Van Der Heide, B., Kim, S., Westerman, D., & Tong, S. T. (2008). The role of friends' appearance and behavior on evaluations of individuals on Facebook: Are we known by the company we keep? *Human Communication Research, 34,* 28–49. doi:10.1111/j.1468-2958.2007.00312.x
Watson, J. B. [1914] (1967). *Behavior: An introduction to comparative psychology.* New York: Holt, Rinehart and Winston.
Webb, E. J., Campbell, D. T., Schwartz, R. D., Sechrest, L., & Grove, J. B. (1981). *Nonreactive measures in the social sciences* (2nd ed.). Boston: Houghton-Mifflin.
Weiner, B. (1975). "On being sane in insane places": A process (attributional) analysis and critique. *Journal of Abnormal Psychology, 84,* 433–441.
Weisz, J. R., Jensen-Doss, A., & Hawley, K. M. (2006). Evidence-based youth psychotherapies versus usual clinical care. *American Psychologist, 61,* 671–689.
West, S. G. (2010). Alternatives to randomized experiments. *Current Directions in Psychological Science, 18,* 299–304. doi: 10.1111/j.1467-8721.2009.01656.x
West, S. G., & Thoemmes, F. (2010). Campbell's and Rubin's perspectives on causal inference. *Psychological Methods, 15,* 18–37. doi: 10.1037/a00115917
Whalen, C., & Schreibman, L. (2003). Joint attention training for children with autism using behavior modification procedures. *Journal of Child Psychology and Psychiatry, 44,* 456–468. doi: 10.1111/1469-7610.00135
Whitlock, J. L., Powers, J. L., & Eckenrode, J. (2006). The virtual cutting edge: The Internet and adolescent self-jury. *Developmental Psychology, 42,* 407–417.
Wilkinson, L., & Task Force on Statistical Inference. (1999). Statistical methods in psychology journals. *American Psychologist, 54,* 598–604.
Willis, C. M., Church, S. M., Guest, C. M., Cook, W. A., McCarthy, N., Bransbury, A. J., et al. (2004). Olfactory detection of human bladder cancer by dogs: Proof of principle study. *British Medical Journal, 329,* 712–716.
Willis, J., & Todorov, A. (2006). First impressions: Making up your mind after a 100-ms exposure to a face. *Psychological Science, 17,* 592–598. doi: 10.1111/j.1467-9280.2006.01750.x
Wilson, G. T. (1978). On the much discussed nature of the term "behavior therapy." *Behavior Therapy, 9,* 89–98.

Wilson, R. E., Gosling, S. D., & Graham, L. T. (2012). A review of Facebook research in the social sciences. *Perspectives on Psychological Science, 7*, 203–220. doi: 10.1177/1745691612442904

Winer, B. J., Brown, D. R., & Michels, K. M. (1991). *Statistical principles in experimental design* (3rd ed.). New York: McGraw-Hill.

Yeaton, W. H., & Sechrest, L. (1986). Use and misuse of no-difference findings in eliminating threats to validity. *Evaluation Review, 10*, 836–852.

Zechmeister, E. B., Chronis, A. M., Cull, W. L., D'Anna, C. A., & Healy, N. A. (1995). Growth of a functionally important lexicon. *Journal of Reading Behavior, 27*, 201–212.

Zechmeister, E. B., & Posavac, E. J. (2003). *Data analysis and interpretation in the behavioral sciences*. Belmont, CA: Wadsworth.

Zechmeister, J. S., Zechmeister, E. B., & Shaughnessy, J. J. (2001). *Essentials of research methods in psychology*, New York: McGraw-Hill.

Zimbardo, P. G. (2004). Does psychology make a significant difference in our lives? *American Psychologist, 59*, 339–351.

Grafik, Tablo ve Kutucukların Alındıkları Kaynaklar

Bölüm 1

Şekil 1.1a: © Imagery Majestic/Cutcaster RF; Şekil 1.1b: © Bananastock RF; Box 1.1: © Courtesy of Princeton University; Şekil 1.2a: © Kim Steele/Getty Images RF; Şekil 1.2b and 1.2c: © Corbis RF.

Bölüm 2

Şekil 2.1 (üst ve alt): Courtesy of Thomas A. Sebeok, Distinguished Professor Emeritus, Indiana University, Bloomington; Kutu 2.1: © J. S. Zechmeister; Şekil 2.2: © Paul Bradbury/Getty Images RF; Şekil 2.3 (left): © The Museum of Questionable Medical Devices, www.museumofquackery.com; Şekil 2.3 (Sağ): © Corbis RF; Figure 2.4: © David Buffington/Getty Images RF; Şekil 2.5: Veriler, Levine, R. V., Norenzayan, A., & Philbrick, K. Cross-cultural differences in helping strangers. Journal of Cross-Cultural Psychology, 32, 543-560, (2001) s. 551'deki Tablo 2'den, yazarların izniyle uyarlanmıştır. Copyright © 2001 Sage Publications, Inc.

Bölüm 3

Şekil 3.1 (sol): © Dynamic Graphics/JupiterImages RF; Şekil 3.1 (sağ) © E. B. Zechmeister, Linda & Hilary Bryant, Candy Bauilat ve Peter Berquist'e teşekkür eder; Şekil 3.2 (sol ve sağ): © E. B. Zechmeister (sol) ve © J. J. Shaughnessy (sağ) Bu bölümdeki bazı fotoğrafları temin eden Parmly Institute of Loyola University of Chicago'dan Bill Shofner ve Rich Bowen'e teşekkür ederiz. Şekil 3.3: © Greg Gibson/AP Photo/Wide World Photos; Şekil 3.4: © Digital Vision RF; Şekil 3.5: Photo of Eugene and Jeanne Zechmeister by a friendly passerby; Şekil 3.6: Copyright © 1968 by Stanley Milgram, copyright renewed 1993 by Alexandra Milgram. Pennsylvania State Media Sales'ın dağıtımını yaptığı Obedience filminden alınmıştır.

Bölüm 4

Şekil 4.2: © Brand X/Getty Images RF; Şekil 4.3: © Ira E. Hyman, Jr., Western Washington University; Şekil 4.4: © Brand X Pictures/Punchstock RF; Table 4.3: Ölçek, Hope College, Psikoloji bölümünden Jane Dickie'den sağlanmıştır. MI; Tablo 4.4: Nairan Ramirez-Esparza ile Ramirez-Esparza, N., Mehl, M. R. Alvarez-Bermudez, J., & Pennebaker, J. W. Are Mexicans more or less sociable than Americans? İnsights from a naturalistic study. Journal of Research in Personality, 43, 1-7 s. 5'teki verilerden alınmıştır. Şekil 4.5: © Jim Sugar/Corbis.

Bölüm 5

Şekil 5.2: © Bananastock RF; Şekil 5.3: Sax, L. J., Austin, A. W., Lindholm, J. A., Korn, W. S., Saenz, V. B., & MAhoney, K. M. (2003). The

American Freshman: National norms for fall 2003. Los Angeles: Higher Education Research Institute, UCLA, s. 7'deki Şekil 7'den alınmıştır. *The American freshman: National norms for fall 2003.* Los Angeles: Higher Education Research Institute, UCLA. © UC Requests; Şekil 5.4: © Duncan Smith/Getty Images RF; Şekil 5.5: © Ryan McVay/Getty Images RF; Table 5.1: Lucas, R. E., Diener, E., & Suh, E. Discriminant validity of well-being measures. Journal of Personality and Social Psychology, 71, 616-628 Copyright, 1996, American Psychological Association, sayfa 621 Tablo 3'teki verilerden, yazarların izniyle uyarlanmıştır. Şekil 5.6: © Ingram Publishing/AGE Fotostock.

Bölüm 6

Şekil 6.1: © Indiapicture/Alamy Images; v 6.2: Henny Ray Abrams/AFP/Getty Images; Şekil 6.3: © Corbis RF; Tablo 6.1: Carnagey, N. L., & Anderson, C. A. The effects of reward and punishment in violent video games on agressive affect, cognition, and behavior. Psychological Science, 16, 882-889. Copyright 2005 Association for Psychological Science sayfa 885, Tablo 2'den, yayımcıların izniyle uyarlanmıştır; Şekil 6.4: © Brand X Pictures RF; Figure 6.5: © Photodisc/Getty Images RF.

Bölüm 7

Şekil 7.1: © Ryan McVay/Getty Images RF; Şekiller 7.2a ve 7.2b: © Image Source/Getty Images RF; Şekil 7.3: © Comstock/PunchStock RF.

Bölüm 8

Tablo 8.3: Veriler, Saul M. Kassin'den alınmıştır. Şekil 8.1: Kassin, S. M., Goldstein, C. C., & Savitsky, K. Behavioral confirmation in the interrogation room: On the dangers of presuming guilt. Law and Human Behavior, 27, 187-203. Copyright 2003 American Psychological-Law Society/Division 41 - American Psychological Association, s. 198, Şekil 2'den uyarlanmıştır. Veriler, Saul M. Kassin'den alınmıştır.; Şekil 8.5: Pingitore, R., Dugoni, B. L., Tindale, R. S. & Spring, B. Bias against overweight job applicants in a simulated employment interview. Journal of Applied Psychology, 79, 909-917. Copyright 1994 American Psychological Association s. 913, TAblo 2'den uyarlanmıştır.; Tablo 8.5 ve 8.6 ve Şekil 8.6: Kaiser, C. R., Vick, S. B., & Major, B. Prejudice expectations moderate preconscious attention to cues that are threatening to social idendity. Psychological Science, 17, 332-338. Copyright 2006 Association for Psychological Science s. 336'da sunulan verilerden uyarlanmıştır.

Bölüm 9

Şekil 9.1 (sol ve sağ): © Stacy M. Lopresti-Goodman, Department of Psychology, Marymount University, Arlington, VA; Box 9.2: Vaka çalışması örneği 1978 Division of Psychotherapy (29), American Psychological Association'dan yayımcıların ve yazarların izniyle uyarlanmıştır. Bu materyal referans verilirken gösterilmesi gereken kaynak: Kirsch, I. (1978). Teaching clients to be their own therapists: A case study illustration.

Psychotherapy: Yheory, Research and Practice, 15, 302-305. Bu bilginin kullanılması, yayımcıların onayladığı anlamına gelmez.; Şekil 9.2: © LWA-Dann Tardif/Corbis; Şekil 9.3: Horton, S. V. Reduction of disruptive mealtime behavior by facial screening. Behavior modification, 11, 53-64. Copyright 1987 Sage Publications, Inc s. 60, Şekil 1'den yayımcı ve yazarın izniyle uyarlanmıştır. Şekil 9.4: Bu şekil, Lang, R., Regester, A., Mulloy, A., Rispoli, M., & Botout, A. Behavioral intervention to treat selective mutism across multiple social situations and community settings. Journal of Applied Behavioral Analysis, 44, 623-628. Copyright 2011 Society for the Experimental Analysis of Behavior s. 627, Şekil 1'den yazarın izniyle uyarlanmıştır..

Bölüm 10

Şekil 10.1: © Children's Television Workshop/Hulton Archive/Getty Images; Şekil 10.2: © Mikael Karlsson/Arresting Images RF; Şekil 10.3: © Ryan McVay/Getty Images RF; Şekil 10.5: Peterson, C., & Seligman, M. E. P. Character strengths before and after September 11. Psychological Science, 14, 381–384. Copyright © 2003 Association for Psychological Science, sayfa 383, Şekil 1'den uyarlanmıştır; Şekil 10.6 ve 10.7: Khuder, S. A., Milz, S., Jordan, T., Price, J., Silvestri, K., & Butler, P. The impact of a smoking ban on hospital admissions for coronary heart disease. Preventive Medicine, 45, 3–8. Copyright © 2007 Elsevier s. 5, Şekil 1 ve 3'teki verilerden uyarlanmıştır.; Şekil 10.8: Salomon, G. Basic and applied research in psychology: Reciprocity between two worlds. International Journal of Psychology, 22, 441–446. Copyright © 1987 International Union of Psychological Science and Psychology Press, s. 444, Şekil 2'den alınmıştır.

Bölüm 11

Şekil 11.2a ve 11.2b: Fotoğraflar, J. S. DeLoache, et al., 1998, "Grasping the Nature of Pictures," in Psychological Science, 9, 205–210 Şekil 2'den alınmıştır; Şekil 11.3 ve 11.4: Judy DeLoache'den alınmıştır ve DeLoache, J. S., Pierroutsakos, S. L., Uttal, D. H., Rosengren, K. S., & Gottlieb, A. Grasping at the nature of pictures. Psychological Science, 9, 205–210. Copyright © 1998 Association for Psychological Science Şekil 3'ten yayımcı ve yazarın izniyle uyarlanmıştır.

Bölüm 13

Alıntılar, Publication Manual of the American Psychological Association, 6th Edition (2010), Washington, DC: Copyright © 2010 American Psychological Association sayfa 65 (Writing Styles), sayfa 70-71 (Reducing Bias in Language), açımlanmış metin veya içerik sayfa 71-77 (General Guidelines for Reducing Bias) ve Bölüm 2 (MAnuscript Structure and Content), bölümler (Abstracts, Introduction, Method, Results, and Discussion) ve ss. 25-26'dan alınmıştır. Bu materyale referans verilirken kullanılması gereken resmî alıntı:

American Psychological Association (2010). *Publication manual* (6th ed.). Washington, DC: Author.

İsim Dizini

Abelson, R. P., 195, 343–345, 361, 380, 385, 388
Adler, T., 13
Ajarova, L., 279
Akugizibwe, T., 279
Albarracin, D., 29
Allen, H. L., 308
Allport, G. W., 43, 284, 303
Almeroth, K., 57
Álvarez-Bermùdez, J., 122, 168
Ambady, N., 38, 45
American Psychiatric Association, 40–42
American Psychological Association (APA), 9, 11, 13–14, 23, 59, 85, 382, 418
Anand, V., 371
Ancell, M., 108
Anderson, C. A., 4, 196–199, 208
Anderson, C. R., 66
Anderson, J. R., 50
Anderson, K. J., 269
Anglin, J. M., 346
Aristotle, 6
Arkowitz, H., 17
Armeli, S., 118
Armstrong-Stasson, M., 66
Arnett, J. J., 11
Association for Psychological Science (APS), 9, 11, 15, 59, 149, 418
Astin, A. W., 152–153
Astone, N. M., 159
Atkinson, R. C., 283
Atwater, J. D., 334–335

Back, M. D., 66, 112
Baer, D. M., 299
Baicker, K., 308
Baker, T. B., 15, 118
Baldwin, A. K., 147
Banaji, M. R., 29, 60, 64, 66, 78, 128, 148–150, 208
Bard, K. A., 96
Barlow, D. H., 286, 294, 297
Barnett, D., 10
Baron, R. M., 170
Bartholomew, G. A., 418
Bartlett, M. Y., 34
Bass, K. E., 18
Bathje, G. F., 229
Baugh, D., 99
Bauhaus, S., 96
Bauman, C. W., 66
Baumeister, R. F., 20, 93, 168–169
Baumrind, D., 73, 76
Bauserman, R., 10
Baxter, M. G., 154–157
Becker-Blease, K. A., 64
Begley, S., xi
Behnke, S., 75, 82, 84
Bellack, A. S., 67, 77, 78
Bentham, J., 81
Berdan, J., 153–154
Bergsieker, H. B., 12
Berk, R. A., 334
Berkowitz, L., 4
Berliner, L., 10
Bernstein, M., 308
Bimber, B., 57
Birnbaum, M., 10–11, 148–150
Blake, L. P., 153–154
Blanchard, F. A., 128
Blanck, P. D., 67, 77, 78
Blass, T., 75, 84
Bleich, A., 66
Boekhoven, B., 96
Bolgar, H., 286
Bonanno, G. A., 66
Boring, E. G., 30
Boruch, R. F., 334
Boss, S. M., 103–104, 114–115
Botout, A., 294–295
Bower, G. H., 267–269
Bowker, A., 96
Brandt, R. M., 113
Bransbury, A. J., 32
Brechan, I., 29
Brigham, J. C., 128
Britton, B., 77
Broffman, M., 32
Brossart, D. F., 288
Brotsky, S. R., 103
Brown, D. R., 233
Brown, R., 50
Bruckman, A., 60, 64, 66, 69, 78, 128, 148–150
Buchanan, T., 60, 64, 128, 149
Buhrmester, M., 149
Bulger, M., 57
Bullock, J. J., 159
Burger, J. M., xiii, 76, 84, 89
Burr, J. E., 99
Bushman, B. J., 187, 208, 218
Butler, P., xiv, 328, 330–331

Caggiano, J. M., 103–104, 114–115
Campbell, D. T., 53, 107, 109–111, 126, 310, 313, 314, 315–316, 318, 322, 323, 324, 326, 328–329, 334, 338, 385
Campbell, J., 57
Candland, D. K., 100
Carey, B., 284
Carnagey, N. L., 196–199
Case, M. H., 153–154
Ceci, S. J., 206
Cepeda-Benito, A., 10
Chabris, C. F., 17, 18, 29
Chaffin, M., 10
Chambers, D. L., 334
Chapman, G. B., 163
Charness, N., 277
Chastain, G., 61
Chernoff, N. N., 8
Chow, S. L., 380, 390
Christensen, L., 73, 76
Chronis, A. M., 346

Chun, D., 57
Chung, J., 334
Church, S. M., 32
Cicchetti, D., 104–105, 113, 130
Clark, C. S., 95, 128
Clark, M. S., 137
Cohen, A. B., 11
Cohen, J., 60, 64, 66, 69, 78, 128, 148–150, 199–200, 344, 354–355, 361, 380, 384, 385, 388, 391, 398–399
Congressional Budget Office, 335
Connor, N., 277
Conway, A. R. A., 51
Cook, T. D., 53, 310, 314, 315–316, 322, 323, 324, 326, 328–329, 338, 385
Cook, W. A., 32
Coon, D. J., 6
Cordaro, L., 130
Cordon, I., 10
Corkin, S., 284
Corrigan, P., 229
Costall, A., 96
Couper, M., 60, 64, 66, 69, 78, 128, 148–150
Craik, K. H., 7
Crandall, C. S., 128
Crandall, R., 59, 70, 128
Creswell, J. W., 119
Crick, N. R., 99
Cronbach, L. J., 421
Crook, M. D., 18
Crossen, C., 137, 144
Crossley, A. M., 169
Crowder, R. G., 208
Cull, W. L., 346
Cullerton-Sen, C., 99
Cumming, G., 360, 365–366, 380
Curtiss, S. R., 100

Dalal, R., 11, 149
Dallam, S. J., 10
D'Anna, C. A., 346
Darley, J. M., 168
Darwin, 7
Dawes, R. M., 206
DeAngelo, L., 152–153
DeLeeuw, K., 57
DeLoache, J. S., 362–366
Descartes, 6
DeSteno, D., 34
Deters, F. G., 95
Dickie, J. R., 115–117
Diener, E., 19, 59, 70, 128, 161–162, 170
Dittmar, H., 25, 182–187
Dixit, J., 103, 128
Dolan, C. A., 117–118
Donnerstein, E., 4, 190
Downs, M. F., 229
Drabick, D. A., 99
Dube, A., 278–279
Dubner, S. J., 111
Duckett, E., 95
Dugoni, B. L., 253–255
Dunlosky, J., 3
Durham, D. L., 279
Durtschi, J., 137

Eagan, K., 153–154
Eagly, A. H., 29
Eastwick, P. W., 3
Ebbinghaus, H., 277
Ebneter, D., 229
Eckenrode, J., 100, 120, 121, 128
Egloof, B., 66, 112
Eibl-Eibesfeldt, I., 98
Einstein, A., 7
Eisenberg, D., 229
Ekman, P., 93
Elms, A. C., 75
Eloul, L., 12
Endersby, J. W., 108
Entwisle, D. R., 159
Epley, N., 76
Epstein, S., 161
Ericsson, K. A., 277
Estes, W. K., 359, 366, 380
Evans, A. D., 105
Evans, G. W., 170–171
Evans, R., 190
Eysenbach, G., 100

Fechner, G. T., 277
Feeney, D. M., 80
Ferdowsian, H. R., 279
Fernandez, K. C., 110, 124
Field, A. E., 154–157
Finch, S., 360, 365–366, 380
Fincham, F. D., 137
Fine, M. A., 82, 89
Finkel, E. J., 3
Finkelstein, A. N., 308
Fiore, M. C., 118
Fischler, C., 108, 110
Fisher, C. B., 63
Fleischman, D. A., 277
Foa, R., 137
Fossey, D., 127
Fowler, R. D., 80
Fox, J., 151
Fraley, R. C., 148
Francis, M. E., 21, 180
Freeman, P. R., 237
Freud, S., 8, 40
Freyd, J. J., 64
Frick, R. W., 385
Fryberg, D., 63
Funder, D. C., 20, 93, 168–169

Gabrieli, J. D. E., 277
Gaddis, S., 112
Galileo, 7
Galvani, A. P., 163
Geier, A., 108
Gelkopf, M., 66
Gena, A., 296
Gentile, L., 170–171
Gerrard, M., 334–335
Gigerenzer, G., 25
Giles, D., 103
Gilman, R., 277
Glaser, J., 103, 128
Gleaves, D. H., 10
Glover, P., 96
Golberstein, E., 229
Goldstein, C. C., 246–251, 264
Gonczy, C., 32
Gonnella, C., 170–171
Gonzales, R., 380, 381
Goodall, J., 80
Goodman, G. S., 10
Gordon, R. T., 32
Gosling, S. D., 7, 11, 93, 107, 110, 112, 148, 149
Gottlieb, A., 362–366
Graham, L. T., 110
Graham, S. M., 137

Green, D. P., 103, 128
Greenfield, P. M., 100, 113, 119
Greenwald, A. G., 29, 60, 64, 78, 380, 381
Griffiths, M. D., 137
Griskevicius, V., 38–39
Grissom, R. J., 354, 361
Grove, J. B., 107, 109–111, 126
Gruber, J. H., 308
Guest, C. M., 32
Günaydin, G., xiii, 221
Guthrie, D., 380, 381

Hagen, R. L., 380
Haggbloom, S. J., 8
Halliwell, E., 25, 182–187
Halpern, A. R., 267–269
Hamedani, M. G., 12
Haney, M., 277
Harlow, H. F., 100
Harlow, M. K., 100
Harré, R., 278, 284
Harris, R. J., 380–381
Harsch, N., 51
Harshman, R. A., 359, 380, 385, 388
Hart, W., 29
Hartup, W. W., 99
Hawkins, D. L., 99
Hawley, K. M., 200
Hazan, C., xiii, 221
Healy, N. A., 346
Heatherton, T. F., 137, 154–157
Hemmerich, J. A., 51
Hersen, M., 286, 294, 297
Hilts, P. J., 283
Hippler, H. J., 147
Hoaglin, D. C., 344, 347, 348
Hodgins, D. C., 229, 238
H. M. (see Molaison, H. G.)
Holden, C., 62, 80
Holland, R. W., 25
Hollerran, S. E., 95, 125, 126, 128
Holmbeck, G. N., 95, 170
Holsti, O. R., 119
Horch, J. D., 229, 238
Horton, S. V., 292–293
Hou, M., 206
Howell, D. C., 359, 373, 380, 385
Hubbard, A., 32
Huberman, A. M., 120
Huesmann, L. R., 4
Huff, C., 76
Hume, D., 81
Hunt, M., 199
Hunter, J. E., 344, 380, 385, 388
Hurtado, S., 152–153
Husain, G., 18
Hussain, Z., 137
Hsu, L. M., 334
Hyman, I. E., 103–104, 114–115

Inglehart, R., 37
Ison, J. R., 130
Itakura, S., 96
Ivceric, Z., 38
Ive, S., 25, 182–187

James, W., 8
Janecki, T., 32
Jansen-Yeh, E., 99
Jensen-Doss, A., 200
Jezierski, T., 32
John, O. P., 11, 149
Johnson, C. M., 279
Johnson, D., 79
Johnson, G. M., 334–335
Johnson, J. A., 148
Johnson, J. D., 4
Joiner, T. E., 154–157
Jones, D., xiii, 11
Jones, N. A., 159
Jones, V. K., 8
Jordan, T., xiv, 328, 330–331
Judd, C. M., 123, 199

Kabnick, K., 108, 110
Kahneman, D., 8, 29, 312
Kaiser, C. R., 257–262, 411
Kameka, M., 278–279
Kandel, E., 284
Karney, B. R., 3
Kaschak, M. P., 347
Kasser, T., 13–14
Kassin, S. M., 75, 77, 83, 246–251, 264
Kawakami, K., 25
Kazdin, A. E., xiv, 280, 282, 290, 293–294, 297–299
Keane, M. M., 277
Keating, C. F., 99
Keel, P., 154–157
Kelman, H. C., 69, 76
Kenny, D. A., 170
Keppel, G., 248, 384, 408, 412
Kershaw, T. C., 51
Khuder, J. A., xiv, 328, 330–331
Kidd, S.A., 44, 119
Kidder, L. H., 123, 199
Kiechel, K. L., 75, 77, 83
Kim, J. J., 354, 361
Kim, S., 112
Kimble, G. A., 40, 53
Kimmel, A. J., 70, 73, 74, 76
Kimwele, C., 279
Kirigin-Ramp, K., 334–335
Kirk, R. E., 354–355, 398
Kirkham, G. L., 102
Kirsch, I., 194, 280–282
Klaiman, C., 277
Klein, E., 307, 335
Klinesmith, J., 13–14
Knafo, A., xiii, 46
Knight, A., 57
Korn, W. S., 152–153
Kosty, M., 32
Kraemer, H. C., 10
Kral, M. J., 44, 119
Kramer, A. D. I., 10, 110
Kranendonk, G., 279
Krantz, J. H., 11, 149
Krantz, P. J., 296
Kratochwill, T. R., 288–289
Kraut, R., 60, 64, 66, 69, 78, 128, 148–150
Krendl, A. C., 38
Kroenër, J., 32
Krueger, J., 380
Kruger, J., 16
Kubany, E. S., 277
Kufner, A. C. P., 66, 112
Kulik, J., 50
Kurdek, L. A., 82, 89
Kwang, T., 149
Ky, K. N., 18

Lambert, N. M., 137
Landrum, R. E., 61
Lang, R. R., 294–295
Langer, E. J., xiv, 320–326, 334
Larson, R. W., 95, 118, 125
Latané, B., 168
Latner, J. D., 229
LeBlanc, P., 62
Lee, K., 96, 105
Lee, L., 147
Lee, S. W. S., 206
Levin, J. R., 288
Levine, R. V., xiii, 42–43, 46
Levinson, C. A., 110, 124
Levitt, S. D., 111
Li, M., 163
Light, K. C., 117–118
Lilienfeld, S. O., 14, 15–16, 17
Lindberg, M. J., 29
Lindholm, J. A., 152–153
Linz, D., 4
Locke, 6
Locke, T. P., 334–335
Loftus, E. F., 10
Loftus, G. R., 344, 365, 380, 406
Lopresti-Goodman, S. M., xiv, 278–279
Lorenz, K., 8
Lucas, R. E., 161–162

MacCoun, R., 8
Madon, S., 271
Mahamedi, F., 154–157
Mahoney, K. M., 152–153
Major, B., 257–262, 411
Malamuth, N. M., 4
Marcynyszyn, L. A., 170–171
Markus, H. R., 11, 12
Marsh, E. J., 3
Martens, B. K., 289
Marx, M. H., 28, 30, 33, 39, 51
Masson, M. E. J., 365, 406
Masters, J. C., 286
Matsumoto, D., 99
Matsuzawa, T., 96
Maugh, T. H., 284
Mayer, R. E., 57
Mazur, E., 120
McAndrew, F. T., 13–14
McCallum, D. M., 62
McCarthy, A., 96
McCarthy, D. E., 118
McCarthy, N., 32
McClannahan, L. E., 296
McCoy, T. P., 308
McCulloch, M., 32
McFall, R. M., 15
McGuire, W. J., 21
McKenzie, K. E., 103–104, 114–115
McRaney, D., 29
Meehl, P. E., 53, 380
Mehl, M. R., xiii, 95, 122, 125, 126, 128, 168
Merrill, L., 29
Michels, K. M., 233
Michielutte, R., 329
Miles, M. B., 120
Milgram, S., xiii, 73–75, 84, 89
Mill, J. S., 81
Miller, D., 16
Miller, G. A., 346
Miller, N. E., 79
Milner, B., 283
Milz, S., xiv, 328, 330–331
Molaison, H. G., 284
Moneta, G., 95
Mooallem, J., 98
Mook, D. G., 205, 208
Moore, C. F., 347
Moore, K. M., 51
Morrell, F., 277
Mosteller, F., 344, 347, 348
Muir, D. W., 96
Mulaik, S. A., 359, 380, 385, 388
Mulcahy, J. B., 279
Mullen, E., 66
Mullet, N., 308
Mulloy, A., 294–295
Myers, D. G., 170
Myers, L. J., 32
Myowa-Yamakoshi, M., 96

Nathan, M. J., 3
National Academy of Sciences, 79, 80
Neisser, U., 7, 51
Nettle, D., 106
Newburger, E. C., 149
Newhagen, J. E., 108
Newhouse, J. P., 308
Newman, E., 62
New York Times Book Review, 107
Nichols, P., 156
Nolan, A., 96
Norenzayon, A., 42–43, 46
Nosek, B. A., 60, 64, 78
Novak, M. A., 79

O'Brien, M. C., 308
Office for Human Research Protections, 71
Olson, J., 60, 64, 66, 69, 78, 128, 148–150
Ondersma, S. J., 10
Orne, M. T., 78, 125, 194
Ostrov, J. M., 99
Otali, E., 279

Park, C. L., 118
Park, N., 34
Parker, R. I., 288
Parnell, P., 98
Parry, H. J., 169
Parsons, H. M., 315
Parsonson, B. S., 299
Pashler, H., 196
Paskett, E. D., 329
Patuki Blake, L., 152–153
Paty, J., 118
Pavlov, I., 8
Pease. A., 93
Pease, B., 93
Pegalis, L., 77
Pennebaker, J. W., xiii, 21, 95, 122, 168, 179–180
Pepler, D. J., 99
Pete, E., 108, 110
Peterson, C., xiv, 34, 137, 327, 329
Philbrick, K., 42–43, 46
Piaget, J., 105
Piasecki, T. M., 118
Pickren, W. E., 8
Pierroutsakos, S. L., 362–366
Pingitore, R., 253–255
Piper, A. I., 334
Pollick, A., 62, 64
Popper, K. R., 50, 53

Posavac, E. J., 167, 332, 347, 361, 366, 382, 386, 390, 402, 408
Poulson, C. L., 296
Poulton, E. C., 237–238
Powell, K. L., 106
Powell, T., 96
Powers, J. L., 100, 120, 121, 128
Price, J., xiv, 328, 330–331
Price, M., 13
Pryor, J. H., 152–153, 229

Raju, N. S., 359, 380, 385, 389
Ralston, P., 99
Ramírez-Esparza, N., xiii, 95, 122, 168
Rasinski, K. A., 147
Rauscher, F. H., 18
Rawson, K. A., 3
Regester, A., 294–295
Reis, H. T., 3
Reminger, S. L., 277
Revelle, W., 269
Richards, M. H., 95
Richardson, D. R., 77
Richardson, J., 98
Riley, D. A., 206
Rimm, D. C., 286
Rind, B., 10
Rispoli, M., 294–295
Robbins, M. L., 95
Roberts, G., 106
Robins, R. W., 7
Rodebaugh, T. L., 110, 124
Rodin, J., xiv, 320–326
Roethlisberger, F. J., 315
Rogers, A., 183
Rogosch, F. A., 104–105, 113, 130
Rosengren, K. S., 362–366
Rosenhan, D. L., 101–102, 129
Rosenthal, R., 45, 63, 130, 194, 355, 361, 388, 398, 410
Rosnow, R. L., 67, 77, 78, 361, 388, 398, 410
Rossi, P. H., 334
Rotheram-Borus, M. J., 67, 77, 78
Rozin, P., 20, 108, 110
Rule, N. O., 38
Russell, T. M., 8

Sacks, O., 277–278, 283
Saenz, V. B., 152–153
Saldaña, J., 120
Salomon, G., 333–334
Salpekar, N., 170–171
Sapirstein, G., 194
Sargent, J. D., 137
Sargis, E. G., 11, 60, 74, 76, 128, 149–150
Savitsky, K., 246–251, 264
Sax, L. J., 152–153
Schacter, D. L., 283
Schatz, C. B., 32
Schaus, J. F., 308
Schellenberg, E. G., 18
Scherr, K. C., 271
Schlessinger, L., 10
Schmidt, F. L., 344, 380, 385, 386, 388
Schmidt, W. C., 149
Schmukle, S. C., 112
Schooler, N. R., 67, 77, 78
Schreibman, L., 296
Schultz, R. T., 277
Schwartz, P., 136
Schwartz, R. D., 107, 109–111, 126
Schwartz, S. H., xiii, 46
Schwarz, N., 147, 206
Scoville, W. B., 283
Sechrest, L., 107, 109–111, 126, 385
Selcuk, E., xiii, 221
Seligman, M. E. P., xiv, 34, 327, 329
Shadish, W. R., 53, 310, 315–316, 322, 324, 326, 328–329, 385
Shapiro, K. J., 79
Shaughnessy, J. J., 41, 51, 83, 209, 238, 267, 319
Shaw, G. L., 18
Shelton, B., 329
Shepperd, J. A., 68
Sherwood, A., 117–118
Shields, C., 108, 110
Shiffman, S., 118
Shiffrin, R. M., 283
Shoham, V., 15
Silberg, J. L., 10
Silver, N., 29
Silverman, D., 44, 119–120
Silvestri, K., xiv, 328, 330–331
Simon, H. A., 8, 50
Simons, D., 17, 29
Singer, P., 81
Skinner, B. F., 9, 288
Skitka, L. J., 11, 51, 60, 66, 74, 76, 128, 149–150
Slatcher, R. B., 95, 99
Smith, E. R., 123, 199
Smith, J. A., 278, 284
Smith, T. W., 166
Smith, V., 8
Snibbe, A. C., 11
Sokal, M. M., 6
Sole, M. L., 308
Solomon, Z., 66
Sperry, R. W., 8
Spiegel, D., 10
Spitz, R. A., 100
Spitzer, R. L., 101
Sprecher, S., 3
Spring, B., 253–255
Srivastava, S., 11, 149
Stanley, J. C., 318, 322, 329
Stark, D. E., 277
Steele, K. M., 18
Steen, T. A., 34
Stephens, N. M., 12
Sternberg, R. J., 20, 204
Stopfer, J. M., 112
Striepe, M., 154–157
Stull, A., 57
Subrahmanyam, K., 100, 113, 119
Sue, S., 206
Suh, E., 161–162
Surwit, R. S., 79
Susskind, J. E., 29
Swami, V., 108

Talarico, J. M., 51
Tanaka, M., 96
Tatum, C. M., 329
Taubman, D., 206
Taubman, S. L., 308
Taylor, K. M., 68
Taylor, S., 96
Tennen, H., 118

Thioux, M., 277
Thoemmes, F., 307
Thomas, L., 180
Thomason, N., 380
Thompson, W. F., 18
Till, J. E., 100
Tinbergen, N., 8
Tindale, R. S., 253–255
Todorov, A., 224–227
Tomonaga, M., 96
Tong, S. T., 112
Toth, S. L., 104–105, 113, 130
Towle, M. J., 108
Tran, S., 152–153
Trentacosta, C. J., 99
Tromovitch, P., 10
Tukey, J. W., 344, 348–349, 380
Turner, K., 32
Tuten, T. L., 149
Tversky, A., 8, 29
Tybur, J. M., 38–39
Tynes, B., 100, 113, 119

Ulrich, R. E., 79–80
Underwood, B. J., 209, 238, 267
Uttal, D. H., 362–366
Utz, S., 103, 113

Valentino, K., 104–105, 113, 130
van Baaren, R. B., 25
Van den Bergh, B., 38–39
Van Der Heide, B., 112
VanderStoep, S. W., 319
van Knippenberg, A., 25
Van Langenhove, L., 278, 284
Vaughn, L. A., 128
Vazire, S., 11, 95, 112, 128, 149
Velez, R., 329
Vick, S. B., 257–262, 411
Vietri, J., 163
Vohs, K. D., 20, 93, 168–169
von Békésy, G., 8

Wagenmakers, E. J., 196
Wakefield, P. C., 346
Walther, J. B., 112
Wansink, B., 108
Warber, K. M., 151
Warnick, J. E., 8
Warnick, R., 8
Wartella, E., 4
Watson, J. B., 7, 30
Webb, E. J., 107, 109–111, 126
Weiner, B., 101
Weisz, J. R., 200
Welzel, C., 137
West, S. G., 307, 310
Westerman, D., 112
Whalen, C., 296
Whitlock, J. L., 100, 120, 121, 128
Wilkinson, L., 389
Williams, J. E., 60, 64, 128
Williams, P. G., 79
Willingham, B., 99
Willingham, D. T., 3
Willis, C. M., 32
Willis, G. B., 147
Willis, J., 32, 224–227
Wilson, G. T., 288
Wilson, R. E., 110
Winer, B. J., 233
Wirtz, D., 16
Wise, B. M., 103–104, 114–115
Witte, A. D., 334
Wundt, W., 6

Yarbrough, G. L., 8
Yeaton, W. H., 385
Yeh, W., 147

Zaayer, J., 32
Zayas, V., xiii, 221
Zechmeister, E. B., 41, 51, 83, 167, 346, 347, 361, 366, 386, 402, 408
Zechmeister, J. S., 41, 51, 83
Zhang, H., 57
Zimbardo, P. G., 49
Zivin, K., 229

Konu Dizini

ABAB (tersine) desen, 291-294
  etik meseleler, 294
  yöntemsel meseleler, 292-294
ABBA karşıt dengeleme, (tekrarlı ölçümler desenleri), 226-228
Akıllı Hans, 30-32
Akran değerlendirmesi, 418
Alan deneyleri, 105-106, 206-209
Alan notları, 113
Aldatıcı korelasyon, 29
Alfa. *Bk.* Anlamlılık düzeyi
Alışma (uyum), 126
Alışma etkileri (tekrarlı ölçümler desenleri), 222-223
Alışma etkilerinin aşamaları. *Bk.* Alışma etkileri
Amerikan Psikiyatri Birliği, Amerikan Psikologlar Birliği (APA), 40-42
  Hayvan Araştırmaları Etik Komitesi (CARE), 79
  Etik Yönetmeliği, 13-14, 59-86
  tarihi, 9
  APA Yayım Kılavuzu, 85, 386, 390, 403, 419-420, 422-423, 432-434
  web sitesi, 9, 59, 85, 419-420, 432-433
Anket
  cevaplarının analizi,167
  doğruluğu ve dakikliği, 158-160
  geçerliği, 160-162
  güvenirliği, 160-162
  hazırlanmasında aşamalar, 162-163
  soruların etkili ifadesi (rehber), 163-166
  soruların ifadesi, 163-166
  soruların sırası, 166-167
  tanımı, 158
  tercihler ve tutumlar, 160
  ve demografik değişkenler, 158-159
  ve öz bildirim, 160-162
Anket araştırmaları
  araç olarak anketler, 158-160
  ardışık bağımsız örneklemler deseni, 152-154
  boylamsal araştırma, 154-157
  etik meseleler, 137, 150
  geçerliği, 160-162
  güvenirliği, 160-162
  hata payı, 357
  internet araştırmaları, 148-150
  kesitsel desen, 151-152
  kişisel görüşmeler, 146-147
  kullanımı, 136-137
  özellikleri, 138
  postayla araştırma, 145-146
  seçkisiz rakam çevirme, 147-148
  telefon görüşmeleri, 147-148
  ve anket hazırlama, 162-167
  ve korelasyonel araştırma, 136, 167, 169-171
  ve örneklem seçme, 138-144
  ve sosyal istenirlik, 168
  ve tepkisellik, 168-169
  yanlılık, 140, 145, 146
Anket araştırmalarında örnekleme çerçevesi, 139
Anlamlılık düzeyi, 202, 381. *Ayrıca bk.* Sıfır Hipotezi Anlamlılık Testi (SHAT)
Anlamlılık, 388. *Ayrıca bk.* İstatistiksel anlamlılık
Anlatımsal kayıt
  alan notları, 113
  analizi, 118-120
  kodlanması, 119
  tanım, 113
  veri indirgeme, 119
ANOVA. *Bk.* Varyans analizi (ANOVA; *F* testi) Beklenti etkisi, 228
APA *Yayını El Kitabı*. *Bk.* Amerikan Psikologlar Birliği (APA)
Aracı değişken, 170-171
Aracı değişkenler, 52-53
Araç düzeni (iç geçerliğe tehdit), 310-311
Araçlar, 36-38. *Ayrıca bk.* Derecelendirme ölçekleri; Anketler
Aralık, 352. *Ayrıca bk.* Betimsel istatistikler; Değişkenlik (ölçümler)
Araştırma (süreç)
  araştırmaya başlama, 18
  aşamaları, 22
  bir araştırmacı gibi düşünmek, 14-17
  çok yöntemli yaklaşım, 23, 179
  niteliksel ve niceliksel, 44
  uygulamalı ve temel, 49, 305, 333-336
Araştırma bulgularının genelliği, 42, 48-49, 284. *Ayrıca bk.* Dış geçerlik

Araştırma desenleri/yöntemleri (tipleri)
alan deneyleri, 105-106, 206-209
anket araştırması, 144-157
bağımsız gruplar, 182
deneysel, 31-33, 179-181
dikkat çekmeyen (tepkisel olmayan), 106-112
doğal gruplar, 211-213, 266-269
eşlenmiş gruplar, 209-211
gözleme dayalı, 97-106
karışık, 245, 411-412
karmaşık, 244
korelasyonel, 136, 167, 169-171
psikofizik, 222
seçkisiz gruplar, 181-187
tek denekli desen (küçük-*n*) deneysel, 289-297
tekrarlı ölçümler, 220, 224
vaka araştırması, 278
yarı deneysel yöntemler, 317-331
Araştırma önerileri, 435-436
Araştırma raporu yazma, 418-436.
American Psychological Association (APA)
analiz süreci, 344-345
APA *Yayım Kılavuzu*.
Ayrıca bk. Sonuçların bildirilmesi
etik meseleler, 80-85, 434
etkili yazma kılavuzu, 421-423
raporun yapısı, 423-432
referanslar (atıf yapma), 432-433
yazma ipuçları, 419, 424-429, 432
Araştırma sonrası katılımcıları bilgilendirme, 14*n*, 76-78
Ardışık bağımsız örneklemler deseni (anket araştırması), 152-154
Ardışık kayıtlar, (Arşiv kayıtları), 110
Aritmetik ortalama (averaj). *Bk.* Ortalama
Arşiv kayıtları (veriler)
içerik analizi, 119-120
kodlama, 119
olay kayıtları, 110
seçici birikim, 111-112
seçici kalma, 112
sorunları ve sınırlılıkları, 111-112
sürekli kayıt, 110
tanımı, 110
tipleri, 107, 110
ve doğal müdahale, 110-111
ve kesintili zaman dizileri deseni, 326
ve mantığı, 110
ve program değerlendirme, 332-333
ve sahte ilişki, 112
ve vaka çalışmaları, 278
Artık değişkenlik, 403-406
Aşırmacılık, 82-83
Averaj. *Bk.* Ortalama
Ayıt edici geçerlilik, 162
Ayrıca bk. Sonuçların bildirilmesi; Araştırma Raporunun yazılması
Bağımlı değişken (tanımı), 33, 181. *Ayrıca bk.* Davranışın ölçülmesi
Bağımsız değişken
ilişkili olan/olmayan, 265
karıştırıcı, 48,185, 309-314
seçilmiş, 32*n*, 211
tanımı, 32
ve bireysel farklar, 211
ve doğal gruplar, 211-213
ve tek denekli (küçük-*n*) ve deney, 289-291
Bağımsız değişkenin değişimlenmesi, 31-32, 183-185, 187, 306-307.
Ayrıca bk. Deney
Bağımsız gruplar deseni (şekilleri)
bütün hâlinde gruplar, 189-190, 313
denk gruplar, 211-213
doğal gruplar, 211-213
iki değişkenli, 244-245, 407
karışık gruplar, 411-412
karmaşık, 244
seçkisiz gruplar, 181-187
tanımı, 182
tek değişkenli, 182, 393
üç değişkenli, 253-255
Basit kesintili zaman dizileri deseni, 326-330
Basit seçkisiz örnekleme (seçkisiz örnekleme), 142
Basit temel etki, 259, 408. *Ayrıca bk.* Temel etki
Başlama sırasının seçkisiz başlaması (karşıt dengeleme), 232
Bekleme listesi kontrol grubu, 309
Beklenti etkileri (deneycinin/gözlemcinin), 129-130, 194
Betimsel istatistikler, 121-123. *Ayrıca bk.* Merkezî eğilim; Yayılım ölçüleri
Bilgilendirilmiş onam, 67-76. *Ayrıca bk.*Psikoloji araştırmalarında etik meseleler
Bilimde sezginin rolü, 29-30
Bilimsel dergiler (psikolojide). *Ayrıca bk.*
alıntı yapma, 82-83, 425, 432-433

arama, 426
elektronik dergiler, 421
kabul (ret) oranı, 418
Psikolojide İletişim
Bilimsel psikoloji
ve kanıtlar, 3-9, 14-18
ahlâkî bağlam, 13-14
ve Nobel Ödülü, 8
tarihsel bağlam, 6-9
sosyokültürel bağlam (Zeitgeist), 9-12
Bilimsel yöntem
amaçları, 40-49
araçlar, 36-37
bağlamı, 6-12
bilimsel olmayan ve bilimsel yaklaşımlar, 28
hipotez testi, 19, 38-40
niteliksel ve niceliksel analiz, 44
özellikleri, 3-5, 14-18, 28-54
tanım, 4-5
teori oluşturma ve test etme, 50-54
ve doğrulama yanlılığı, 29
ve görgül yaklaşım, 5-9
ve iletişim, 17-18, 35-36
ve kavramlar, 33-35
ve klinik psikoloji, 15
ve kontrol, 30-33
ve sezgi, 29
Bilimsel yöntemin amacı olarak betimleme, 40-41
Bilimsel yöntemin amacı olarak değişiklik yaratma. *Bk.* Bilimsel yöntemin amacı olarak uygulama.
Bilimsel yöntemin amacı olarak kestirim, 44-46
Bilimsel yöntemin amacı olarak uygulama, 40, 49, 331-336
Bilimsel yöntemin bir amacı olarak açıklama, 46-49, 179. *Ayrıca bk.* Nedensel çıkarım; Deneysel yöntemler
Bilimsel yöntemin bir amacı olarak anlama. *Bk.* Bilimsel yöntemin bir amacı olarak açıklama
Bilişsel psikoloji, 7
Bireysel farklar (denek) değişkeni
tanımı, 32n, 211
ve doğal gruplar deseni, 211-213
ve tekrarlı ölçümler desenleri, 222-223
Blok seçkisizleştirme, 188-189, 224-226
Boylamsal desen (anket araştırmaları), 154-157
Bütün hâlinde gruplar, test etkisi, 189-190, 313. *Ayrıca bk.* İç geçerliğe tehdit
CARE. *Bk.* Hayvan Araştırmaları ve Etik Komitesi
Cohen's *d*, 199, 354-355, 387, 402
tanımı, 199, 354
Cohen's *f*, 399
Çıkarımsal istatistik, 201. *Ayrıca bk.* Sıfır Hipotezi Anlamlılık Testi (SHAT)
Çıkarma yöntemi, 253
Çift amaçlı soru (anket), 164-165
Çift kör işlemi, 194-195
Çizgi grafikler, 253, 432
Çok yöntemli yaklaşım (hipotez testine), 23, 179
Çoklu temel düzey deseni
bireyler arasında, 296
davranışlar arasında, 296
durumlar arasında, 294
yöntemsel meseleler, 297
Çubuk grafikler, 253, 431
d (etki büyüklüğü). *Bk.* Cohen *d*.
D çubuğu (farklı puanların ortalamaları), 361-368
Davranış değişimi, 288
Davranış tedavisi, 288-289
Davranışçılık, 7, 9
Davranışın deneysel analizi, 278, 288
Davranışın kaydı
alan notları, 113
amaçları, 112
anlatı kayıtları, 113
elektronik kayıtlar, 95, 117-118
göreceli sıklık, 121
gözlemsel yöntemlerin sınıflanması, 97
kapsamlı kayıtlar, 112-114
seçilmiş kayıtlar, 112-118
sıklığı, 114, 117, 290
Davranışın ölçülmesi. *Ayrıca bk.* Davranışın kaydedilmesi
davranış ölçmede anketler, 158
davranış ölçüm araçları, 36-37
davranış ölçümünde değişkenlik, 122, 352-354
davranış ölçümünde doğruluk, 36
davranış ölçümünde duyarlılık, 36
davranış ölçümünde geçerlik, 38
davranış ölçümünün güvenirliği, 38, 123-124
davranışın fiziksel/psikolojik ölçümü, 37-38
derecelemeler, 115-117
dikkat çekmeyen (tepkisel olmayan) davranış ölçümü, 106-112

elektronik davranış ölçümü, 95, 99, 117-118
göreceli sıklık, 121
kapsamlı/seçilmiş yönlerinin ölçülmesi, 112-118
nitel ve nicel davranış ölçümü, 44, 118-123
ölçekler, 115-117
öz bildirim, 160
ve bilimsel yöntem, 36-38
ve tepkisellik, 100, 106-107
Değişkenler (tipleri)
aracı, 52-53
bağımlı, 32, 181
bağımsız, 32, 181
bireysel farklar, 32*n*, 211-213
değişimlenmiş, 31-32, 183-185
demografik, 158-159
dış, 190-191
doğal gruplar. *Bk.* Bireysel farklar
eşleştirme, 210
ilişkili bağımsız, 264-265
ilişkisiz bağımsız, 264-265
katılımcı. *Bk.ç* Bireysel farklılıklar
seçilmiş, 32*n*
yönlendirici, 170-171
Değişkenlik (ölçümler). *Ayrıca bk.* Hata değişkenliği; Yayılım ölçüleri,
aralık, 352
standart sapma, 121, 352-353
Demografik değişkenler, 158-159
Denek değişkeni. *Bk.* Bireysel farklar değişkeni
Denek içi desen. *Bk.* Tekrarlı ölçümler Deseni
Denekler arası desen. *Bk.* Bağımsız gruplar deseni,
Deney,
alan deneyleri, 105-106, 206-209
analizi, 195-205, 346-373, 380-413
deneyin mantığı, 180-181
deneylerin duyarlılığı, 221, 384
dış geçerliği, 205-209
gerçek deneyler, 186-187, 306-307
iç geçerliği, 47-48, 189, 309-314
laboratuvarda/doğal ortamda, 305-306
tanımı, 31-33, 46-47
tek denekli deneyler (küçük-*n*), 288-300
toplum deneyleri, 334-336
ve nedensel çıkarım, 46-49179-180
yarı deneysel yöntemler, 306, 316-331
Deneycinin (gözlemcinin) etkileri, 106-107, 125-127, 129-130, 194-195. *Ayrıca bk.* Tepkisellik
tanımı, 194
Deneysel gruplar arasında iletişim (iç geçerliğe yönelik tehdit), 314
Deneysel koşul (müdahale), 31-33, 186-187
Dengeleme. *Bk.* kontrol teknikleri
Denk olmayan kontrol grup deseni
iç geçerliğe tehditler, 321-325
örneği, 319-325
tanımı, 318-319
ve dış geçerlik, 325-326
Denk olmayan kontrol gruplu zaman dizileri deseni, 301-331
Dereceleme ölçekleri, 115-117. *Ayrıca bk.* Davranışın ölçümü
Dış değişkenler, 190-191
Dış geçerlik
dış geçerliğe tehditler, 305, 316
dış geçerliğin sağlanması, 205-209
dış geçerlik deneyleri, 205-209, 316
tanımı, 94, 205
ve etkileşim, 263-265
ve laboratuvar araştırmaları, 305
ve örnekleme, 94-97
ve replikasyonu, 205-209, 316
ve tek denekli (küçük-*n*) deneysel desenler, 299-300
ve vaka araştırması, 287
Diferansiyel aktarma sorunu, 237-238
Diferansiyel istatistiksel regresyon, 324. *Ayrıca bk.* Seçimin ilave etkileri; İç geçerliğe tehdit
Dikkat çekmeyen gözlem (tepkisel olmayan)
arşiv kayıtları, 107, 109-112
etik meseleler, 129
fiziksel izler, 107-109
geçerliği, 108-109, 111-112
gözlemler
kullanım izleri, 107
tanım, 106
ürünler, 108
ve tepkisellik, 106
Doğal gözlem, 99-100
Doğal gruplar değişkeni. *Bk.* Bireysel farklar değişkeni
Doğal gruplar deseni, tanımı, 212

ve bireysel farklar değişkeni, 211
ve etkileşim etkisi, 266-269
ve karmaşık desen, 266-269
ve nedensel çıkarım, 212-213
Doğal izlerin kullanımı, 107-108
Doğal müdahale, 110-111, 327
Doğrulama yanlılığı, 29
Doğrusal eğilim, 370. *Ayrıca bk.* Korelasyon
Dolaylı gözlem. *Bk.* Gözlem; Dikkat çekmeyen ölçümler
Durum örnekleme, 96
Duyarlılık (deneysel), 221, 384-386
Duyarsızlaşma (adaptasyon), 127
Elektronik dergiler, 421
Elektronik olarak harekete geçen kayıt cihazı, 95
Eşit aralıklı ölçek, 115-117
Eşlenmiş gruplar deseni, 209-211
Eta kare (η^2), 199, 235, 398, 410-412
tanımı, 398
Etik uygunluk kontrol listesi, 85
Etki büyüklüğü (ölçümleri)
Cohen's *d*, 199, 354-355, 387, 402
Cohen's *f*, 399
eta kare (η^2), 235, 398, 410, 412
iki faktörlü bağımsız gruplar deseni, 410
sonuçların bildirilmesi, 355, 390, 413, 429
tanımı, 198, 354
ve bağımsız gruplar desenleri, 198-199, 387, 398-399
ve güç analizi, 391-392, 399-400
ve meta-analiz, 199-200
ve tekrarlı ölçümler desenleri, 235-236, 388
Etkileşim, 244, 247, 250, 255. *Ayrıca bk.* Varyans analizi (ANOVA); Karmaşık desenler
betimleme, 251-255
iki yönlü (faktörlü), 250-253
tanımı, 250
üç yönlü (faktörlü), 254-255
ve çıkarma yöntemi, 253
ve dış geçerlik, 263-265
ve ilişkili değişkenler, 264-265
ve tavan (taban) etkileri, 265-266
ve teori test etme, 262-263, 266-269
yorumlama, 262-266
Etkili yazma rehberi, 421-423. *Ayrıca bk.* Araştırma raporu yazma
Etnoloji, 8, 98
Evren
örneklem ve evren, 140, 144, 202
parametre, 356
tanımı, 139
ve örneklem seçme, 140-144
f (etki büyüklüğü). *Bk.* Cohen's *f*
Facebook, 110-112
Faktöryel kombinasyon (desen), 244, *Ayrıca bk.* karmaşık desenler
Farklı Puanlar, 360-362
Faydacılık, 81
F-dağılımı, kritik değerler, 440-441
Filtre sorular (ankette), 166
Fiziksel izler (dikkat çekmeyen ölçümler), 107-109
F-testi (tanımı), 393-394. *Ayrıca bk.* Varyans analizi (ANOVA)
Geçerlik (tipleri)
ayırt edici, 162
dış, 94, 205-209
iç, 181-189
ortak noktada birleşme, 161, 179
tanım, 38
yapı, 161-162
Genel *F*-testi, 394. *Ayrıca bk.* Varyans analizi (ANOVA)
Gerçek deney
kontrol edilemeyen sorunlar, 314-316
özellikleri, 307
ve iç geçerliğe tehditler, 309-313
yapılmasındaki engeller, 309
Gizlilik, 64, 66-67, 70-73
Göreceli sıklık, 121
Görgül yaklaşım, 5-9
tanımı, 7
Görüşme. *Bk.* Kişisel görüşmeler
Görüşmeci yanlılığı, 146-147
Gözlem
alan deneyleri, 105-106
bilimsel ve bilimsel olmayan gözlem, 28, 29
dikkat çekmeyen (tepkisel olmayan) gözlem,112-116
doğal gözlem, 99
doğrudan ve dolaylı gözlem, 97, 106
etik meseleler, 103, 127-129
gözlemcinin etkisi, 101-102, 106-107, 125-127
güvenirliği, 123-124

kapsamlı ve seçici kayıtlar, 112
katılımcılar, 101-102
körleme gözlem, 130, 194-195
ve elektronik aygıtlar, 95, 117-118, 122
ve internet, 102-103
ve kontrol, 30-33
ve müdahale, 100-106
ve örneklem seçme, 94-97
ve talep özellikleri, 125
ve tepkisellik, 101, 106-107, 125-127
yanlılığı, 129-130
yapılandırılmış gözlem, 103-105
yöntemleri ve sınıflaması, 97
Gözlemci (deneyci) yanlılığı, 129-130
Gözlemci güvenirliği
tanımı, 123
ve bilimsel raporlar, 35-36
ve ölçümü, 123-124
Gözlemciler arası güvenirlik, 123-124.
Ayrıca bk. Gözlemci güvenirliği
Gözlemcilerin uyum oranı, 123
Grafikler (çizme)
çizgi grafik, 253, 431
çubuk grafik, 253, 431
sütun grafik, 430-431
Grup yöntemi, 42, 277, 284.
Ayrıca bk. Nomotetik yaklaşım
Güç (istatistiksel)
etkileyen güçler, 384-385
tanımı, 384
ve deneysel duyarlılık, 384-386
ve örneklem büyüklüğü, 384
Güven aralığı, 390
bir evrendeğer için, 356
ikiden fazla bağımsız ortalama, 362-366
rapor edilmesi, 390, 428-429
tanımı, 203-204, 356, 359, 361
ve bağımsız gruplar deseni, 203-204, 359-361, 362-363
ve etki büyüklüğü, 361
ve iki bağımsız grubun ortalamaları, 359-361
ve istatistiksel anlamlılık, 203-204, 365-366
ve karma desenler, 410-411
ve korelasyon, 373
ve tek ortalama, 356-359
ve tekrarlı ölçümler deseninde iki ortalama, 360-362
ve tekrarlı ölçümler desenleri, 360-362
yorumu, 203-204, 359, 361, 365-366
Güvenirlik
deneysel güvenirlik, 38, 196
gözlemci güvenirliği, 123-124
ölçümü, 38
tanımı, 38
test-tekrar test, 160
ve öz bildirim ölçümleri, 160-161
ve tekrar, 196
H_0. *Bk.* Sıfır hipotezi (H_0)
Haber olma uğursuzluğu, xiv, 312
Hata çubukları, 357, 365-366
Hata değişkenliği, 201-202, 235-237, 396-397, 405.
Ayrıca bk. Karelerin ortalama hatası
Hata sınırları, 357
Hawthorne etkisi, 315-316.
Ayrıca bk. Yenilik etkileri
Hayvan Araştırmaları ve Etik Komitesi (CARE), 79
Hipotez(ler). *Ayrıca bk.* Sıfır Hipotezi Anlamlılık Testi (SHAT)
geliştirme, 18-21
ve çok yöntemli yaklaşım, 23, 179
sıfır hipotezi, 202,381, 394.
ve bilimsel teoriler, 39
test edilebilirlik, 40
tanımı, 21,38
Huni soru (anket), 166
IRB. *Bk.* Kurumsal İnceleme Kurulu
İç geçerliğe denek seçimi tehdidi. *Bk.* İç geçerliğe seçim tehdidi
İç geçerliğe tehdit olarak bozucu etki, 315
İç geçerliğe tehditler. *Ayrıca bk.*
araçlandırma, 310-311
bütün hâlinde gruplar (test etme), 189-190, 313
diferansiyel istatistiksel regresyon, 324
kayıp, 191-194, 312
kirlenme, 314
olgunlaşma, 310
regresyon (istatistiksel), 311
seçim, 312-313
seçimin ilave etkisi, 303
tanımı, 189, 309-310
tarih, 310
test etme, 310
ve alışma etkileri, 222-224
ve gerçek deneyler, 309-314
yenilik etkisi, 315
yerel tarih etkisi, 323-324

zaman-sıra ilişkisi (ve nedensel çıkarım), 47, 185, 187
İç geçerlik. *Ayrıca bk.* Kontrol teknikleri
deneylerde, 47-48, 189, 223-224,309-314
iç geçerliğe tehdit, 189-195, 223-224, 309-313
plasebo kontrolü, 194-195
tanımı, 181
ve bütün hâlinde (intact) gruplar, 189-190, 313
ve çift körleme işlemi, 194-195
ve dış değişkenler, 190-191
ve gerçek/yarı deneysel yöntemler, 309-313
yarı deneysel desenler, 316, 325
İçerik analizi, 119
İdiografik yaklaşım, 29, 43-44, 284. *Ayrıca bk.* Vaka araştırması yöntemi.
İki ortalamanın karşılaştırılması, 199, 359-362, 386-392, 401-403
İlişkili bağımsız değişken, 264-265
İlişkisiz bağımsız değişken, 265
İnternet günlüğü, 118
İnternette araştırma
elektronik dergiler, 421
internet tartışma grupları, 420
internette anket, 148-151
internette etik meseleler, 11, 60, 69, 71, 76, 78, 137, 150
internette katılımcı kaydı, 10-11, 149-150
internette katılımcı siteleri, 11
internette literatür araştırma, 426
internetten alıntı yapma 432-433
ve bilimsel psikoloji, 10-11
ve davranış kaydı, 110, 112, 119
ve gözlem, 100, 102-103, 110, 112
veri tabanları, 420-421
İnternette araştırma. *Bk.* İnternet
İstatistiksel anlamlılık testi. *Bk.* Sıfır Hipotezi Anlamlılık Testi (SHAT). *Ayrıca bk. t*-testi ve *F*-testi
İstatistiksel anlamlılık. *Ayrıca bk.* Sıfır Hipotezi Anlamlılık Testi (SHAT)
testleri, 202-203, 380-383. *Ayrıca bk. t*-testi ve *F*-testi
ve alfa, 202, 381
ve bilimsel ve pratik anlamlılık, 388-389
ve güç, 384-386
İstatistiksel gerileme. *Bk.* Regresyon (istatistiksel)
İstatistiksel olarak anlamlı (tanım), 202
İstatistiksel testler. *Bk.* Çıkarımsal İstatistikler; Sıfır hipotezi anlamlılık testi (SHAT); *Ayrıca bk. t*-testi ve *F*-testi
İstisnaî puanlar. *Bk.* Uç puanlar
İşbirlikçi, 103
İşevuruk tanım
tanımı, 34
ve eleştiriler, 34-35
ve iletişim, 35
İz çözümlemesi, 170
Kapalı soru (anket), 164
Karma desen, 245, 411-412
Karıştırıcı. *Ayrıca bk.* Makul alternatif nedenler; İç geçerliğe tehditler
tanımı, 48, 185
ve dış değişkenler, 190-191
ve iç geçerlik, 181, 189-195, 309-314
Karmaşık (faktöryel) desenler, 243-273
analizi, 255-262, 407-413
betimleme etkileri, 244-245
faktöryel desen, 244
faktöryel kombinasyon, 244
iki faktörlü, 245-253, 407-412
karma desen, 245, 411-412
sonuçlarını bildirme, 413
tanımı, 244
üç faktörlü, 253-255
ve basit temel etki, 259, 408
ve etkileşim etkisi, 244, 257-260, 407-409
ve temel etki, 244, 247-249, 253-255, 260-262, 409-410
ve teori test etme, 262-263, 266-269
Karşıt dengeleme, 224-233. *Ayrıca bk.* Kontrol teknikleri
Kategorik ölçek, 115-116, 167
Kategorik veri. *Bk.* Sıralı ölçek
Katılımcı gözlemi, 101-103
Katılımcıları aldatma, 14, 73-76, 84
Kavram, *Ayrıca bk.* Yapı.
Kavramsal tekrar (replikasyon), 208
Kayıp, 191-104, 312
Kesintili zaman dizileri deseni
basit şekli, 326

denk olmayan kontrol gruplu şekli, 330
dış geçerlik, 330
iç geçerlik, 329-331
Kesitsel araştırmalar (anket araştırmaları), 151-152
Kestirim (istatistiksel). *Bk.* Korelasyon
Kısmî tekrar, 208
Ki-kare olasılık testi, 167
Kirlenme (iç geçerliğe tehdit), 314
Kişisel görüşme, 146-147
Kodlama,
soruları cevaplama, 167
tanımı, 119
ve içerik analizi, 119-120
ve verilerin indirgenmesi, 119
Kontrol koşulu, 32, 181, 194
Kontrol listesi, 117
Kontrol teknikleri
ABBA karşıt dengeleme, 228-233
blok seçkisizleştirme, 188-189, 224-226
çift-körleme işlemi, 194-195
dengeleme, 185-187, 201, 228-233
karşıt dengeleme (tanımı), 224
koşulları sabit tutma, 185
Lâtin karesi, 232-233
plasebo kontrolü, 194-195
seçkisiz atama, 182, 185-187, 201
seçkisiz başlayan dönüşümlü sıra, 232
ve değişimleme, 31-33, 183-185
ve tek denekli (küçük-*n*) deney, 289-291
Kontrol, *Ayrıca bk.* Kontrol koşulu; Kontrol teknikleri
Kontrollü kullanım (planlı kullanım) izleri, 108
Korelasyon
güven aralığı, 373
katsayısı, 124, 136, 371-372
negatif, 124, 371
Pearson-Çarpım Momentleri korelasyonu, 123-124, 372
Pozitif, 124, 371
tanımı, 45, 124, 367
ve bilimsel yöntem, 40-41, 44-46
ve doğrusal eğilim, 370
ve gözlemci güvenirliği, 123-124
ve nedensellik, 44-46, 169-171, 372
ve prediksiyon, 44-46, 136
ve saçılım grafiği, 369-371
ve sahte ilişki, 170
yanıltıcı, 29
Korelasyon katsayısı. *Bk.* Korelasyon
Korelasyonel araştırma, 136. *Ayrıca bk.* Korelasyon
Koşulları sabit tutma. *Bk.* Kontrol teknikleri
Kök ve gövde gösterimi, 348-350
Kör (deneyci, gözlemci), 130, 194-195
Kötü deney. *Bk.* Tek grup ön test-son test deseni
Kritik değer. *Bk.* İstatistiksel anlamlılık.
Kullanım izleri (dikkat çekmeyen ölçümler), 107
Kurumsal Hayvan Bakım ve Kullanım Kurulu (IACUC), 60-62
Kurumsal İnceleme Kurulu (IRB), 60-62
Kültür merkezli yanlılık, 11-12
Kültürler arası araştırma, xiii, 11-12, 42-43, 96, 108-109, 122, 362-366
ve etnik merkezci yanlılık, 11-12
ve internet araştırmaları, 149
Latin karesi (karşıt dengeleme), 232-233
Makul alternatif nedenler (karıştırıcılar), 47-48, 309-310. *Ayrıca bk.* İç geçerliğe Tehditler
Medyada araştırma bulguları, 17-18
Mekanik denek kaybı, 191
Merkezî eğilim ölçüleri, 121, 351-352, *Ayrıca bk.* Ortalama; Ortanca (Medyan); Mod
Merkezî eğilim, 121, 197-198, 351-352 *Ayrıca bk.* Ortalama; Medyan (Ortanca); Mod
Meta analiz, 199-200, 355, 429
Milgram'ın deneyi, 74-75, 84
Minimal risk, 65
Mod, 351. *Ayrıca bk.* Merkezî eğilim
Müdahale aşamaları ve temel düzey, 290-291
Müdahale. *Bk.* Deneysel durum; Bağımsız değişken; Değişimleme; Doğal müdahale
Mümkün bütün sıralar (karşıt dengeleme), 230-231
N = 1 desenleri. *Bk.* Tek denekli (küçük-*n*) deneysel desenler
Nedensel çıkarım (ilişki) koşulları, 47

olası alternatif nedenler, 47, 181, 187, 309-310
tanımı, 47
ve bilimsel yöntem, 46-49
ve deneyler, 46-47
ve doğal gruplar desenleri, 212-213, 266-269
ve iç geçerlik, 181
ve korelasyon, 47, 169-171, 372
ve vaka araştırmaları, 285-286
Negatif korelasyon, 124, 371. *Ayrıca bk.* Korelasyon
Nicel araştırma. *Ayrıca bk.*
ölçekler, 115-117
tanımı, 44
ve nitel araştırma, 44
veri analizi, 120-124
Niteliksel araştırma
anlatı kayıtları, 113-114
tanımı, 44
ve içerik analizi, 119-120
ve kodlama, 119
ve nicel araştırma, 44
veri indirgeme, 119
verilerin analizi, 118-120
Nobel ödülü, 8
Nomotetik yaklaşım, 42, 277, 284
Olay kaydı (Arşiv verileri), 110
Olay örnekleme, 96
Olgunlaşma (iç geçerliğe tehdit), 310
Oranlı ölçek (ölçüm), 115-117
Ortak noktada birleşen geçerlik, 179. *Ayrıca bk.* Çok yöntemli yaklaşım (hipotez testinde)
Ortalama kareler hatası. *Bk.* Varyans analizi (ANOVA)
Ortalama, 121, 197,-198, 352
Ortalamanın standart hatası, 353-354
Ortalamanın tahminî standart hatası, 353-354. *Ayrıca bk.* Ortalamanın standart hatası
Ortanca (Medyan), 351-352. *Ayrıca bk.* Merkezî Eğilim
Öge (anket araştırmalarında), 139
Ölçekleme, *Ayrıca bk.* Dereceleme ölçekleri Saçılım grafiği, 369. *Ayrıca bk.* Korelasyon
Ölçekler, 115-117. *Ayrıca bk.* Davranışın ölçümü
Ölçmenin dakikliği. *Bk.* Davranışın ölçülmesi
Örneklem (tanım), 140. *Ayrıca bk.* Örneklemler (tipleri); Örneklem seçme
Örneklem büyüklüğü. *Ayrıca bk.* Serbestlik derecesi (sd)
ve güç, 384-385
ve internet, 149
Örneklem seçme
ardışık bağımsız, 152
çerçeve, 139
denk olmayan ardışık örneklemler, 154
durum örnekleme, 96
elementler, 139
katılımcı örnekleme, 97
olay, 96
seçkisiz (basit), 142
seçkisiz tabakalandırılmış, 143
temel terimleri, 138-144
tesadüfî olmayan, 141-144
tesadüfî, 140-143
uygunluk, 144
ve evren, 139
ve internet araştırması, 10-11, 149-150
ve temsil edicilik, 138, 140
yanlı, 140, 144
zaman örnekleme, 94-96
Öz bildirim ölçümleri,, 160. *Ayrıca bk.* Anketler
Parametre, 356-359
Pearson Çarpım Momentleri *Ayrıca bk.* Korelasyon Korelasyon Katsayısı, 124, 136, 371-372.
Plasebo kontrol grubu, 194-195
Postayla anket, 145-157
Pozitif korelasyon, 124, 371. *Ayrıca bk.* Korelasyon
Program değerlendirme, 306, 331-336
Psikofizik yöntemler (psikofizik), 222
Psikoloji araştırmalarında etik meseleler, 13-14, 59-86
aldatma, 14, 73-76, 84
APA Etik Yönetmeliği, 13-14, 59-86
araştırma sonrası bilgilendirme, 14*n*, 76-78
aşırmacılık, 82-83
bilgilendirilmiş onam, 67-76
en az risk, 65
etik uygunluğun aşamaları, 84-85
etik uygunluk kontrol listesi, 85, 435
gizlilik, 64, 66-67, 70-73
Hayvan Araştırmaları Kurulu ve Etik, 79
karar verme, 59-60, 83-85
kurumsal inceleme kurulu, 60-62
risk belirleme, 63-67

risk/fayda oranı, 62-63
ve anketler, 137, 150,
ve dikkat çekmeyen ölçümler, 129
ve gizlenmiş katılımcı gözlemi, 101, 128
ve gözlem, 127-129
ve hayvan araştırmaları, 78-80
ve internet, 11, 59-60, 78
ve sonuçları bildirme, 80-83
ve tek denekli desenler, 294
Psikolojide iletişim, 35-36, 344-345, 418-436.
Psikolojik Bilimler Birliği (APS), 9, 11, 15, 59, 149, 418
web sitesi, 9, 418
PsycINFO, 426
Puanların yayılımı. *Bk.* Yayılım ölçüleri
Rapor yazma. *Bk.* Araştırma raporu yazma,
Regresyon (istatistiksel; iç geçerliğe tehdit), 311-312
diferansiyel istatistiksel regresyon, 324
Risk (katılımcılar için), 63-67
Risk/fayda oranı, 62-63
Ruhsal Bozuklukların Tanısal ve İstatistiksel El Kitabı (DSM-5), 40-42
Sahte ilişki (değişkenler arasında), 112, 169-171
Ayrıca bk. Nedensel ilişki tanımı, 170
Sayısal nesne belirleyici (doi), 433
Seçici katılımcı kaybı, 191
Seçici koruma (arşiv kayıtları), 112
Seçici saklama (arşiv kayıtları), 111
Seçilmiş bağımsız değişken, 32n. *Ayrıca bk.* Bireysel farklar değişkeni
Seçilmiş sıralar (karşıt dengeleme), 232-233
Seçim (iç geçerliğe tehdit), 312-313. *Ayrıca bk.* Seçimle arttırılmış etki
Seçim yanlılığı (örnekleme), 140
Seçimin ilave etkisi (iç geçerliğe tehdit olarak)
diferansiyel istatistiksel regresyon, 324
ve araç düzeni, 313, 324
ve olgunlaşma, 313, 322-323
ve tarih, 313, 323-324
Seçkisiz atama
bütün hâlinde gruplar, 189-190
tanımı, 182
ve blok seçkisizleştirme, 188-189
ve çıkarımsal istatistik, 201, 393
ve dengeleme, 182, 185-187, 201
ve doğal ortamlar, 307-309
ve gerçek deneyler, 307
ve seçkisiz gruplar deseni, 182
ve yarı deneysel yöntemler, 317
Seçkisiz gruplar deseni. *Ayrıca bk.* Bağımsız gruplar desenleri
analizi, 195-204
örnekleri, 182-187
tanım, 181
ve bağımsız gruplar, 182
ve dış geçerlik, 205-209
ve iç geçerlik, 189-195
Seçkisiz numara çevirme, 147-148
Seçkisiz örnekleme (seçim), 142-143. *Ayrıca bk.* Temsil edicilik (örneklemlerin)
Seçkisiz sayılar tablosu, 438
Serbest cevap soruları (anket), 164
Serbestlik derecesi (*sd*)
F istatistiği, 397, 405, 408-409
t istatistiği, 358, 360, 362, 387, 388
tanımı, 396
ve ANOVA özet tablosu, 396, 404, 408-409, 412
ve güç, 391, 399-400
ve güven aralıkları, 358, 360, 362, 400-401, 406, 410-411
ve kritik değerler, 358, 360, 362, 387
SHAT. *Bk.* Sıfır hipotezi anlamlılık testi
Sıfır hipotezi (H_0) (tanımı), 202, 381, 394. *Ayrıca bk.* Sıfır hipotezi anlamlılık testi (SHAT)
tanımı, 381
Sıfır hipotezi anlamlılık testi (SHAT).
anlamlılık düzeyleri, 202, 381
t-testi, *F*-testi
iki ortalamayı kıyaslamak, 386-392, 401-413
ikiden fazla ortalamayı kıyaslamak, 392-413
sonuçlarını bildirme, 390
sonuçlarını yorumlama, 202-205, 381-383, 389
tanımı, 202, 380-383

ve alfa, 202, 381
ve duyarlılık, 384-385
ve etki büyüklüğü, 384-386
ve güç, 384-386, 391-392
ve hata, 383-386
ve istatistiksel anlamlılık, 202, 381, 389
ve kritik değerler, 439-441
Sıfır noktası, 116
Sıklık (davranışın ölçümü), 114, 117, 290
Sıralı ölçek (ölçekler), 115-117
Sistemli değişkenlik, kaynakları, 256. *Ayrıca bk.* Varyans analizi (ANOVA)
Sonuçları rapor etme (ve bilimsel yöntem), 35-36. *Ayrıca bk.* Araştırma raporu yazma
Sonuçların analizi. *Ayrıca bk.* Verilerin Analizi
Sonuçların rapor edilmesi. *Bk.* Araştırma raporunun yazılması
Sosyal istenirlik, 168
Sözlü bildirimler, 160-162
Sözlü sunum, 434-435
Standart sapma. *Ayrıca bk.* Yayılım
ortalamanın tahminî standart hatası, 353-354
ölçüleri
tanımı, 352-353
Susam Sokağı, 306
Süreksizlik, 328
Şekiller. *Bk.* Grafikler
t testi
bağımsız gruplar için, 387
kritik değerler tablosu, 439
tekrarlı ölçümler (denek içi), 387-388
ve iki ortalamanın karşılaştırılması, 202, 386-388
Tabakalandırılmış seçkisiz örnekleme, 143
Taban (tavan) etkisi, 265-266
Talep özellikleri
kontrolü, 126-127
tanımı, 125, 194
ve çift körleme deneyleri, 104-105
ve plasebo kontrol grubu, 194
Tamamlanmamış tekrarlı ölçümler deseni, 228-233
Tamamlanmış tekrarlı ölçümler deseni, 224-229
Tanıklık, değerlendirme, 287
Tarih (iç geçerliğe tehdit), 310
Tavan (ve taban) etkisi, 265-266
Tek değişkenli bağımsız gruplar deseni, 393-398
Tek denek (küçük-*n*) deneyi
ABAB deseni, 291-294
analizi, 288
avantajları, 290
çoklu temel düzey deseni, 294-297
özellikleri, 289-291
sorunları ve sınırlılıkları, 297-300
tanım, 290
temel aşama, 290-291
tersine desen, 291-291
ve çoklu grup desenleri, 277, 284, 290
ve dış geçerlik, 299-300
ve kontrol, 288
Tek denek araştırması. *Ayrıca bk.* Vaka araştırması yöntemi ve Tek denek (küçük-*n*)deneyi
Tek grup ön test son-test deseni, 317-318
Tekrarlama
kavramsal, 208
kısmî şekli, 208
tanımı, 196
ve dış geçerlik, 208, 316
Tekrarlı ölçümler desenleri
analizi, 233-237
ayrımsal transfer sorunu, 237-238
duyarlılığı, 221, 236
hata değişkenliği, 235-237
kullanılma nedenleri, 220-222
tamamlanmamış şekli, 224, 228-233
tamamlanmış şekli, 224-228
tanımı, 220
ve alışma etkileri, 222-233
ve etki büyüklüğü, 235
ve karma desenler, 411-412
ve *t* testi, 387-388
ve tekrarlanmış ölçümler, 222
Telefon anketleri, 147-148
Temel araştırma
tanımı, 49
ve uygulamalı araştırma, 49, 305, 333-336
Temel düzey aşaması, 290-291, 297-299
Temel etki, 244, 247. *Ayrıca bk.* Varyans Analizi (ANOVA); Karmaşık desen
Temsil edicilik (örneklemin)
durum örnekleme, 96-97
seçkisiz örnekleme, 142-143
tanımı, 140
ve anket araştırmaları, 138, 140
ve dış geçerlik, 94
ve olasılığa dayanan örnekleme, 141-143
ve olay örnekleme, 96
ve uygun örnekleme, 144
ve zaman örnekleme, 94-96
Teori
alanı, 50-51

dakikliği, 53
değerlendirme, 53-54
işlevleri, 51
tanımı, 50-51
test etme, 53-54
ve aracı değişkenler, 51-54
ve deneyler, 179-180, 266-269
ve hipotez, 39
ve vaka çalışması, 282-283
yanlışlama, 53
Teorinin alanı, 50-51
Tepki oranı yanlılığı, 145
Tepki yanlılığı, 165
Tepkisel olmayan ölçümler. *Bk.* Dikkat çekmeyen ölçümler
Tepkisellik, 101, 106-107, 125-127. *Ayrıca bk.* Gözlem; Dikkat çekmeyen (tepkisel olmayan) yöntemler
Tersine desen, 291. *Ayrıca bk.* ABAB deseni
Tesadüfî olmayan örneklem, 141-144
Tesadüfî örnekleme, 140-143
tanımı, 141
Test etkisi (iç geçerliğe tehdit), 310
Test-test güvenirliği, 160
Tip I hata, 204-205, 383
Tip II hata, 204-205, 383
Tutumluluk (kuralları), 53
Tutumluluk ilkesi, 53
Uç değer, 347
Uygulamalı araştırma
tanımı, 49, 305
ve temel araştırma, 49, 305, 333-336
Uygulamalı davranış analizi, 288-289
Uygun örnekleme, 144
Uyum (duyarsızlaşma, alışma), 126-127
Ürünler (dikkat çekmeyen ölçümler), 108
Vaka çalışmaları yöntemi
avantajları, 282-284
dezavantajları, 285-287
gösterim, 280-281
özellikleri, 278-280
tanım, 278
ve tanıklık, 287
ve teori, 282-283
Varyans analizi (ANOVA; *F* testi)
basit temel etki, 259, 408
F testi kritik değerler tablosu,
genel *F* testi, 393-398
kareler ortalaması, 397, 405
mantığı, 393
özet tablosu, 396, 404, 412,
sıfır hipotezi, 394-395
tanım, 392-398
temel etki, 244, 247, 261-261, 407
ve bağımsız gruplar desenleri, 392-398, 407-413
ve etki büyüklüğü, 398-399, 402, 406, 410, 412
ve etkileşim etkisi, 407-410
ve güç, 399-400
ve iki değişkenli bağımsız gruplar desenleri, 407-411
ve istatistiksel anlamlılık, 394-397
ve karışık desenler, 411-412
ve karmaşık desenler, 256, 407-413
ve tek değişkenli bağımsız gruplar deseni, 393-398
ve tekrarlı ölçümler desenleri, 236, 401-406
Veri analizi, 118-124, 195-204, 342-416
aşamaları, 196, 343-344, 380
bilgisayar destekli, 345-346
Hipotezi anlamlılık testi (NHST); Ki-kare olasılık testi; Güven aralığı; Korelasyon; Betimsel istatistik; Sıfır
nitel ve nicel, 118-124
t -testi
Veri analizinde aşamalar, 196, 343-344, 380
Veri tabanı (elektronik), 420-421
Veri temizleme, 347-348
Verileri anlamlandırma, 343. *Ayrıca bakınız,* Veri analizinin aşamaları
Verileri özetleme, 343. *Ayrıca bakınız,* Veri analizinde aşamalar
Verilerin anlamının onaylanması, 343.
Verilerin dönüştürülmesi, 348
Verilerin indirgenmesi, 119. *Ayrıca bk.* Kodlama; Betimsel istatistikler
Yanıltıcı sorular (anket) 164-165
Yapı
geçerliği, 161-162
operasyonel tanımı, 34-35
tanımı, 33-34
ve bilimsel yöntem, 33-35,
Yapılandırılmış gözlem, 103-105
Yarı deneysel yöntemler
analizi, 329
denk olmayan kontrol grubu, 318-326
denk olmayan kontrol gruplu zaman dizileri, 330-331
iç geçerliğe tehditler, 309-313

kesintili zaman dizileri (basit), 326-330
tanımı, 317
ve dış geçerlik, 325-326, 330
ve gerçek deneyler, 306-317
ve iç geçerlik, 321-325, 329-331
Yayılım ölçüleri (değişkenlik), 197-198, 352-353. *Ayrıca bk.* Aralık; Standart sapma
Yayım El Kitabı (APA). *Bk.* Amerikan Psikologlar Birliği (APA)
Yayımlama (araştırma bulgularını), 80-85, 418-420, 434
Yazarların belirlenmesi, 81-82, 424
Yenilik etkisi (iç geçerliğe tehditler), 315-316
Yerel tarih etkisi, 323-324
Ayrıca bk. Seçimin ilave etkisi; İç geçerliğe tehditler
Yönlendirici değişken, 170-171
Zaman dizileri. *Bk.* Basit kesintili zaman dizileri deseni
Zaman örnekleme, 94-96
Zeitgeist, 9

İSTANBUL TİCARET ÜNİVERSİTESİ KÜTÜPHANESİ		
Yayının Geri Getirileceği Tarih		
12 Nisan 2018		
04 Temmuz 2018		
08 -11- 2018		
25 -12- 2018		
11 -10- 2019		
31 -10- 2019		
20 -12- 2019		
13 -01- 2020		
04 -11- 2021		
12 Ocak 2022		
01 -02- 2022		
31 -03- 2022		
02 -06- 2022		
28 -06- 2022		
21 -10- 2022		

http://kutuphane.ticaret.edu.tr
e-mail: kutuphane@ticaret.edu.tr

İSTANBUL TİCARET ÜNİVERSİTESİ KÜTÜPHANESİ

PSİKOLOJİDE *Araştırma Yöntemleri*

Çeviri Editörü: Doç. Dr. İlyas Göz

Research Methods in PSYCHOLOGY

Shaughnessy | Zechmeister | Zechmeister

Bildiğimiz kadarıyla, Türkçe'de sadece psikoloji alanına özgü ve tüm araştırma yöntemlerini içeren kapsamlı bir kitap yoktur. Aslında Bilimsel Araştırma Yöntemleri tüm bilim alanları için aynı temel prensipleri, doğru (geçerli ve güvenilir) bilgiye ulaşmanın yöntemlerini içerir. Bu nitelikteki bir bilgi kontrollü koşullar altında elde edilir ve denetlenmeye, yani test edilmeye ve değişmeye açıktır. Bilimsel bilginin bu özelliği, özellikle lisans öğrencilerinin (ve elbette bu yöntemlerin formel bir eğitimini almamış olan insanların) anlamakta zorlandıkları bir özelliktir, çünkü bilimsel bilginin 'kesin' olduğuna dair yaygın bir yanlış anlama vardır. Yine aynı nedenle, kendi alanında çok başarılı bir hukukçu, işadamı, mühendis, politikacı veya hekimin, geçerliği olmayan bir bilgiye değer vermesi mümkündür.

İşte elinizdeki kitabın temel amacı, karşılaşılan herhangi bir bilginin geçerliğini denetlemeyi mümkün kılacak yöntemleri, prensipleri ve kavramları kazandırmaktır.

Psikolojide Araştırma Yöntemleri kitabının diğer araştırma yöntemleri kitaplarından farkı şudur ki söz konusu yöntemleri ve temel prensipleri açıklamak için psikoloji alanında yapılmış olan araştırmaları ve örnekleri kullanmıştır. Dolayısıyla, araştırma yöntemlerini öğrenmek isteyen psikoloji öğrencileri, rehberlik ve psikolojik danışma öğrencileri ile sosyal hizmet uzmanlığı öğrencileri için özel bir kolaylık sağlamaktadır; soyut prensiplerin ve kavramların anlaşılması, psikoloji alanındaki örneklerle kolaylaşmaktadır. Aynı nedenle mevcut kitap, psikoloji ve psikolojiye yakın alanlarda öğrenim gören lisansüstü öğrenciler ile araştırmacılar için de özel bir destek sağlamaktadır.

Nihayet, *Psikolojide Araştırma Yöntemleri* kitabının endüstri ve örgüt psikolojisi ve sayıları çok olduğu için (yaklaşık, tüm psikologların % 40'ı) özellikle klinik psikoloji alanındaki uygulamacılar için önemine de değinmek uygun olur. Elinizdeki kitabın 15'inci sayfasındaki KUTU 1.2, şu soruyla başlamaktadır: "*Klinik psikologlar, hastalarını tedavi ederken psikoloji araştırmalarının en son bulgularını uygularlar mı?*" Önde gelen klinik psikologlar olan Dr. Timoty Baker, Richard McFall ve Varda Shoham yaptıkları geniş analizde (2008) bu soruya sert ve rahatsız edici bir "Hayır!" cevabı vermişlerdir. Araştırmacılara göre, klinik psikologların büyük kısmı bir araştırmanın yöntemini ve bulgularını anlayabilmek için gerekli olan yöntem bilgisinden yoksun

45 001

KOPYA DEĞİL BENİ OKU

TSYM

Mc Graw Hill Education

Okuyucu öneri ve eleştirileri için okuyucu@nobelyayin.com

facebook.com/nobelyayin

twitter.com/nobelyayin

twitter.com/nobelkitap

ISBN 978-605-

nobel

NOBEL AKADEMİK YAYINCILIK EĞİTİM DANIŞMANLIK TİC. LTD ŞTİ.
Ankara Büro: Mithatpaşa Cad. No.: 74/4 Kızılay ANKARA
Tel: +90312 418 20 10 Faks: +90312 418 30 20
Hobyar Mah. Cemal Nadir Sk. Ferah Han No:16/62 Kat:1
Cağaloğlu-Fatih İSTANBUL Tel-Faks: +90212 513 13 04
nobel@nobelyayin.com - www.nobelyayin.com

İstanbul Ticaret Üniversitesi Kütüphanesi

74501