

2012 / Türkçe Versiyon

Sosyal Bilim Araştırması: Prensipler, Metotlar ve Uygulamalar

Yazan:

Anol Bhattacharjee

University of South Florida, abhattach@usf.edu

Çeviren:

Prof.Dr. Ömer Cevdet Bilgin

Atatürk Üniversitesi, Erzurum, ocbilgin@atauni.edu.tr

Follow this and additional works at: http://scholarcommons.usf.edu/oa_textbooks

 Part of the [American Studies Commons](#), [Education Commons](#), [Public Health Commons](#), and the [Social and Behavioral Sciences Commons](#)

Recommended Citation

Bhattacharjee, Anol, "Social Science Research: Principles, Methods, and Practices" (2012). *Textbooks Collection*. 3.
http://scholarcommons.usf.edu/oa_textbooks/3

This Book is brought to you for free and open access by the USF Tampa Library Open Access Collections at Scholar Commons. It has been accepted for inclusion in Textbooks Collection by an authorized administrator of Scholar Commons. For more information, please contact scholarcommons@usf.edu.

SOSYAL BİLİM ARAŞTIRMASI: PRENSİPLER, METOTLAR VE UYGULAMALAR



**YAZAN:
ANOL BHATTACHERJEE**

**ÇEVİREN:
PROF.DR. ÖMER CEVDET BİLGİN**

SOSYAL BİLİM ARAŞTIRMASI:

PRENSİPLER, METOTLAR VE UYGULAMALAR

Yazan:

Anol Bhattacharjee

University of South Florida

Tampa, Florida, USA

abhattach@usf.edu

Çeviren:

Prof.Dr. Ömer Cevdet Bilgin

Atatürk Üniversitesi

Erzurum, TÜRKİYE

ocbilgin@atauni.edu.tr

Second Edition

Copyright © 2012 by Anol Bhattacharjee

Published under the
Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License



Social Science Research: Principles, Methods, and Practices, 2nd edition
By Anol Bhattacharjee
First published 2012

ISBN-13: 978-1475146127
ISBN-10: 1475146124

Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License:

Users are free to use, copy, share, distribute, display, and reference this book under the following conditions:

- **ATTRIBUTION:** Whole or partial use of this book should be attributed (referenced or cited) according to standard academic practices.
- **NON-COMMERCIAL:** This book may not be used for commercial purposes.
- **SHARE ALIKE:** Users may alter, transform, or build upon this book, but must distribute the resulting work under the same or similar license as this one.

For any reuse or distribution, the license terms of this work must be clearly specified. Your fair use and other rights are in no way affected by the above.

Copyright © 2012 by Anol Bhattacharjee

Çevirenin Notu

Türkiye’de sosyal bilim araştırması metodolojisini öğrenmeye-çalışanlar için yararlı olacağına inandığım, başta bilgi üretiminin felsefi temelleri, terminoloji ve araştırma prosesi olmak üzere bilimsel araştırmanın temel hususiyetlerine dair yazılmış bütüncül ve teferruatsız bu açık erişimli ders notunu Türkçeye çevirmiş olmaktan mutluluk duyuyorum. Sosyal bilim alanlarına mensup lisansüstü öğrencilere verdiğim araştırma metotları dersleri için uygun bir ders notu olarak kullanmaya başladıktan dört yıl sonra bu çeviriyi tamamlamış oldum. Yazarın bilginin evrenselleği ve bilgiye ulaşabilmenin bir insan hakkı olduğu görüşüne kalben katıldığımı bildirmek isterim.

Redaksiyon için yeterli zaman ve çaba eksik kaldı, özür dileyerek, müstakbel okuyucuların nazik katkılarıyla beraber ileride bu eksiği de gidermeyi ümit ediyorum.

Prof.Dr. Ömer Cevdet Bilgin

E-posta: ocbilgin@atauni.edu.tr

Önsöz

Bu kitap, doktora ve yüksek lisans öğrencilerine sosyal bilimler, ekonomi, eğitim, halk sağlığı ve ilgili disiplinlerde bilimsel araştırma prosesini takdim etmek için dizayn edildi. Bu kitap benim South Florida Üniversitesinde on yıldan fazla bir süreyle doktora düzeyinde verdiğim Araştırma Metotları dersinde geliştirilen öğretim materyaline dayanmaktadır. Bu kitap doktora ve yüksek lisans öğrencilerini, yeni başlayan araştırmacıları ve araştırma metotları dersi veren öğretim üyelerini hedeflemektedir. Deneyimli araştırmacılar da yine bu kitabı ellerinin altında bir kompakt referans olarak bulundurabilirler.

Potansiyel okuyucunun bu kitap hakkında sorabileceği ilk ve en önemli soru piyasadaki diğer ders kitaplarından farkının ne olduğudur? Şöyle ki dört kilit fark var. Birincisi, diğer ders kitaplarının aksine, bu kitap yalnızca “araştırma metotları” (ampirik veri toplama ve analiz) hakkında değil aynı zamanda başından sonuna kadar “araştırma prosesinin” bütünü hakkındadır. Araştırma metotları bu araştırma prosesinin yalnızca bir safhasıdır ve belki de en kolay ve en strüktürlü olanıdır. Ders kitaplarının pek çoğu araştırma metotlarını derinlemesine içerir fakat mesela bir araştırmacı gibi teorileştirmek ve düşünmek gibi ampirik araştırmanın çoğu zamanki ön-gereksinimleri olan daha zor, daha az strüktürlü ve muhtemelen daha önemli mevzuları dışarıda bırakırlar. Kendi deneyimime göre, pek çok doktora öğrencisi doktora yılları içerisinde araştırma metotlarında hayli yetkinleşiyorlar fakat ilginç ve yararlı araştırma soruları üretmekte veya bilimsel teorileri inşa etmede zorluk çekiyorlar. Bu eksikliği gidermek gayesiyle, “Bir Araştırmacı gibi Düşünmek” ve “Bilimsel Araştırmada Teoriler” gibi yeni başlayan araştırmacılar için esansiyel olan mevzuları bir bölüm yaptım.

İkincisi, kitabın dizaynı kısa, öz ve kompakttır. Bu kitabı yazarken yalnızca esansiyel kavramlara odaklanmaya ve sayfaları öğrencilerin dikkatini daha az münasebetli veya yanal meselelere saptırabilecek şeylerle doldurmamaya karar verdim. Pek çok doktora semineri ilgili disipline ait yeterince okumayı kapsar. Bu kitap o okumaları tamamlamak için dizayn edildi. Tüm önemli kavramlar biraraya getirilerek özetlendi ve öğrencilerin kendi okumalarına ilave bir yük oluşturacak geniş hacimli bir kitaptan kaçınıldı.

Üçüncüsü, bu kitabın download versiyonu ücretsizdir. Sadece mevcut edisyon değil tüm gelecek edisyonlar da. Kitap aynı zamanda Kindle e-Book, Apple iBook ve cüzi bir ücret karşılığı kağıda basılı olarak ta edinilebilecektir. Bunun sebebi yalnızca öğrencilerimin yüksek kitap fiyatlarından sürekli yakınmaları değil, aynı zamanda bilimsel bilginin ücret ve erişilebilirlik gibi bariyerlerle kısıtlanmaması gerektiğine olan inancımdır. Bilimsel ilerleme yalnızca eğer dünyadaki öğrenciler ve akademisyenler bilimin en iyisine külfetsizce ulaşabilirlerse sağlanır. Bu ücretsiz kitap bu yolda benim naçizane çabamdır. Mamafih, ücretsiz demek “düşük kaliteli” manasına gelmemelidir. Mesela hava, su ve gün ışığı gibi yaşamdaki bazı en güzel şeyler ücretsizdir. Pek çok Google kaynağı da ücretsizdir ve Google olmadan bugünün internet çağında nerede olabileceğimizi bir düşünün. Yine bu kitap gibi günümüzde mevcut Linux ve Apache gibi en sofistike yazılım programları da ücretsizdir.

Dördüncüsü, zaman içerisinde bu kitabın yerel-dil versiyonlarını temin etmeyi planlıyorum ve bu çevirilmiş versiyonlar da ücretsiz olacak. Halihazırda, Çince, Korece, Portekizce ve İspanyolca versiyonları için görüşmeler yaptım (ki 2012 de hazır olacaklarını umuyorum) ve ihtiyaca binaen Arapçaya, Almancaya ve başka dillere çevirmek için nitelikli araştırmacılar veya öğretim üyeleri arıyorum. Eğer benim kibımı çevirmeyi düşünürseniz, şunu bilmelisiniz ki yapacağınız çeviri için herhangi bir finansal kazancınız veya imtiyazınız olmayacak çünkü bu kitabın ücretsiz olması gerekiyor. Fakat, sizi memnuniyetle yerel-dil versiyonuna ortak yazar olarak dahil edebilirim.

Bu kitap 16 bölüm halinde 16 haftalık bir yarıyıl için strüktürlendirildi. Mamafih, öğretim üyeleri veya öğretmenler kendi, müfredatlarının ihtiyaç duyduğu şekilde kitabı kullanmak için, yeni mevzular ekleyebilirler, bazı mevzuları çıkartabilirler veya yoğunlaştırabilirler. Mesela, ben Bölüm 13 ü ve Bölüm 15 i kendi dersime dahil etmiyorum çünkü bizim o materyali ve daha fazlasını içeren daha başka istatistik derslerimiz var. Onun yerine, iki haftayı teorilere (Bölüm 3) harcıyorum, bir haftayı akademik dergiler için gözden-geçirmeleri tartışmaya ve yürütmeye (bu kitapta yok) harcıyorum ve bir haftayı da final sınavına ayırıyorum. En azından, daha başka istatistik derslerine sahip olmayan akademik programlar için Bölüm 14 in ve Bölüm 15 in dahil edilmesini tavsiye ederim. Benim doktora programındaki kendi dersimde kullandığım örnek bir izlenceyi ekte bulabilirsiniz.

Son olarak, bilimsel araştırmada ortaya çıkan yeni trendlere dayalı olarak bu kitabı sürekli güncelleştirmeyi planlıyorum. Eğer gelecek edisyonlarda görmek isteyeceğiniz yeni veya ilginç herhangi bir içerik olursa, lütfen bana bildirin. Ben elimden geleni yapacağım. Mevcut içeriğe dair yorumlar, eleştiriler veya düzeltmeler için şimdiden şükranlarımı sunarım.

Anol Bhattacharjee

E-mail: abhattach@usf.edu

İçindekiler

Araştırmaya Giriş

1. Bilim ve Bilimsel Araştırma	1
2. Bir Araştırmacı gibi Düşünmek	9
3. Araştırma Prosesi	17
4. Bilimsel Araştırmada Teoriler	25

Ampirik Araştırmanın Temelleri

5. Araştırma Dizaynı	35
6. Yapıların Ölçümü	43
7. Ölçek güvenilirliği ve Geçerliliği	55
8. Örneklem	65

Veri Toplama

9. Yoklama Araştırması	73
10. Deneysel Araştırma	83
11. Vaka Araştırması	93
12. İnterpretif Araştırma	103

Veri Analizi

13. Nitel Analiz	113
14. Nicel Analiz: Betimsel İstatistikler	119
15. Nicel Analiz: Çıkarımsal İstatistikler	129

Epilog

16. Araştırma Etiği	137
---------------------------	-----

Ek	143
----------	-----

Bölüm 1

Bilim ve Bilimsel Araştırma

Araştırma (*/research*) nedir? Kime sorduğuna bağlı olarak, basit gibi görünen bu soruya muhtemelen çok farklı cevaplar alırsın. Bazı kişiler, satınalmayı arzu ettikleri ürün ve hizmetleri bulmak için, çevrimiçi internet sitelerini rutin olarak araştırdıklarını söylerler. Televizyon haber kanalları mesela gelecek seçimler veya hükümetin fonladığı projeler gibi konularda halkoyu yoklamaları formunda araştırma yürütürler. Lisans öğrencileri proje ödevlerini veya ev ödevlerini yapmak gayesiyle gereksinim duydukları enformasyonu bulmak için interneti araştırırlar. Lisansüstü öğrenciler bir profesörün danışmanlığında bir araştırma projesinde çalışırken, araştırmayı projelerine ilişkin veriyi toplamak ve analiz etmek olarak görebilirler. İşadamları ve yönetici danışmanları mesela tedarik zinciri daralması gibi organizasyonel problemlere çare bulmak veya müşterilerin satınalma biçimlerini tespit etmek için farklı potansiyel çözümleri araştırırlar. Hangisi olursa olsun, yukarıdakilerden hiçbiri “bilimsel araştırma” olarak telakki edilebilmez, eğer ki (1) bir bilim müktesebatına katkıda bulunmuyor ise ve (2) bilimsel metodu takip etmiyor ise. Bu bölümde bu terimlerin ne manaya geldikleri üzerinde duracağız.

Bilim

Bilim nedir? Kimine göre, bilim lisede veya üniversitede okutulan mesela fizik, kimya ve biyoloji gibi zor derslere tekabül eder ve yalnızca zeki öğrencilerin harcıdır. Kimine göre, bilim beyaz önlüklü bilimcilerin özel ekipmanlar kullanarak laboratuvarlarında icra ettikleri bir iştir. Etimolojik olarak, “bilim (*/science*)” kelimesi “bilgi (*/knowledge*)” manasına gelen Latince “*scientia*” kelimesinden türetilmiştir. **Bilim** herhangi bir tahkikat alanında “bilimsel metot” (bilimsel metodu ileride iyice betimleyeceğiz) kullanılmak suretiyle elde edilen sistematik ve organize bir bilgi yığındır. Bilim iki geniş kategoride gruplandırılabilir: doğa bilimi ve sosyal bilim. **Doğa bilimi** doğal olarak meydana gelen nesnelerin veya fenomenlerin bilimidir, mesela ışık, nesneler, madde, yer küre, göksel cisimler veya insan vücudu gibi. Doğa bilimleri daha detaylı olarak fiziksel bilimler, yer küre bilimleri, yaşam bilimleri ve diğerleri olarak sınıflandırılabilir. Fiziksel bilimler mesela fizik (fiziksel nesnelerin bilimi), kimya (maddenin bilimi) ve astronomi (göksel cisimlerin bilimi) gibi disiplinleri kapsar. Yer küre bilimleri mesela jeoloji (yer kürenin bilimi), meteoroloji (havanın bilimi) ve hidroloji (suların bilimi) gibi disiplinleri kapsar. Yaşam bilimleri mesela biyoloji (canlı organizmaların bilimi), tıp (insan vücudunun bilimi) ve botanik (bitkilerin bilimi) gibi disiplinleri kapsar. Aksine, **sosyal bilim** mesela gruplar, şirketler, toplumlar (*/societies*) veya ekonomiler gibi insanların veya insan topluluklarının ve onların bireysel ve kolektif davranışlarının bilimidir. Sosyal bilimler mesela psikoloji (insan davranışlarının bilimi), sosyoloji (sosyal grupların bilimi) ve ekonomi (şirketlerin, pazarların ve ekonomilerin bilimi) gibi disiplinler halinde sınıflandırılabilir.

Doğa bilimleri sosyal bilimlerden birkaç bakımdan farklıdır. Doğa bilimleri çok kesin (*/hassas/precise*), isabetli (*/accurate*), belirlenimci (*/deterministic*) ve bilimsel gözlemleri yapan kişiden bağımsızdır. Mesela, fizikte sesin belli bir ortamdan geçerken hızını veya suyun ışığı kırma endeksini ölçmek gibi bir bilimsel deney her zaman tam olarak aynı sonuçları verecektir, deneyin ne zaman ve nerede yapıldığına veya deneyi yapan kişiye bakılmaksızın. Eğer aynı fizik deneyini icra eden iki öğrenci o fiziksel özelliğin iki farklı değerini elde eder ise, o zaman genellikle bunun manası öğrencilerden birisinin veya her ikisinin de hata yaptığıdır. Mamafih, sosyal bilimler için aynını söyleyebilmeyiz, öyle ki sosyal bilimler daha

az kesin, daha az isabetli ve daha az belirlenimci olma eğilimi gösterirler. Mesela, bir kişinin mutluluğunu, hipotetik bir enstrüman kullanarak ölçersek, şunu bulabiliriz ki aynı kişi farklı günlerde ve bazen aynı günün farklı zamanlarında daha çok mutludur veya daha az mutludur (ya da mutsuzdur). Kişinin mutluluğu o gün aldığı bir habere veya günün önceki saatlerinde yaşadığı olaylara bağlı olarak değişebilir. Dahası, bir kişinin mutluluğunu isabetle ölçmek için tek bir enstrüman veya metre (*/metric*) mevcut değildir. Bu yüzden, bir enstrüman bir kişiyi “daha çok mutlu” olarak kalibre edebilirken, ikinci bir enstrüman aynı kişinin aynı anda “daha az mutlu” olduğunu belirleyebilir. Başka bir deyişle, sosyal bilimlerde *ölçme hatası* ziyadesiyle mevcuttur ve sosyal bilim politikası kararlarında ehemmiyetli belirsizlik ve anlaşma yoksunluğu vardır. Mesela, doğa bilimcileri arasında, ışığın hızına veya dünyanın güneşin etrafında dönme hızına dair çok anlaşmazlığa rastlanmaz, fakat sosyal bilimciler arasında, mesela küresel terörizmin izalesi veya bir ekonomiyi durgunluktan kurtarma gibi bir sosyal problemin nasıl çözüleceğine dair çok anlaşmazlığa rastlanır. Sosyal bilimleri öğrenmeye-çalışan herhangi bir öğrenci bu bilimlerin ayrılmazı olan muğlaklığın, belirsizliğin ve hatanın yüksek düzeylerdeki mevcudiyetini idrak etmeli ve rahatlıkla ele almalıdır, ki bu zaten sosyal objelerin yüksek değişkenliğini yansıtır.

Bilimler, aynı zamanda, maksatlarına göre sınıflandırılabilir. **Temel bilimler** (*/basic sciences*), aynı zamanda saf (*/pure*) bilimler diye adlandırılırlar, şunlardır ki en temel nesneleri ve kuvvetleri, bunlar arasındaki ilişkileri ve bunları idare eden yasaları açıklarlar. Misal olarak, fizik, matematik ve biyoloji verilebilir. **Uygulamalı bilimler** (*/applied sciences*), aynı zamanda pratik (*/practical*) bilimler diye adlandırılırlar, şu bilimlerdir ki temel bilimlerin bilimsel bilgisini bir fiziksel çevrede uygularlar. Mesela, mühendislik uygulamalı bir bilimdir ki mesela daha dayanıklı köprüler veya yakıt-etkin yanmalı motorlar inşa etmek gibi pratik uygulamalar için fizik biliminin ve kimya biliminin yasalarını uygular. Yine tıp bilimi bir uygulamalı bilimdir ki insanların hastalık problemlerini çözmek için biyoloji biliminin yasalarını uygular. Hem temel bilimler hem de uygulamalı bilimler insan gelişimi için gereklidir. Mamafih, uygulamalı bilimler kendi başlarına ayakta duramazlar, fakat kendi ilerlemeleri için onun yerine temel bilimlere dayanırlar. Elbette, endüstri ve özel sektör pratik değerinden dolayı uygulamalı bilimlere daha fazla odaklanma eğilimindeyken, üniversitelerde hem temel bilimler hem de uygulamalı bilimler öğrenilmeye-çalışılır.

Bilimsel Bilgi

Bilimin maksadı bilimsel bilgi yaratmaktır. **Bilimsel bilgi** (*/scientific knowledge*) ilgilenilen bir fenomeni veya davranışı açıklayan yasaların ve teorilerin genelleştirilmiş bir müktesebatıdır ki bilimsel metot kullanılarak edinilir. **Yasalar** (*/laws*) fenomenlerin veya davranışların gözlenen biçimleri, **teoriler** fenomenin veya davranışın altında yatan sistematik açıklamalardır. Mesela, fizik biliminde, Newtoncu hareket yasaları bir nesne sükunet veya hareketlilik halindeyken ne olduğunu (Newton’un birinci yasası), duran bir nesneyi hareket ettirmek veya hareket eden bir nesneyi durdurmak için ne kadar kuvvet gerektiğini (Newton’un ikinci yasası) ve iki nesne çarpıştığında ne olduğunu (Newton’un üçüncü yasası) betimler. Hepsi beraber, bu üç yasa klasik mekaniğin temelini oluşturur – hareket eden nesnelere dair bir teori. Benzerce, optik teorisi ışığın sahip olduğu özellikleri ve farklı ortamlarda nasıl davrandığını açıklar, elektromanyetik teorisi elektriğin sahip olduğu özellikleri ve nasıl üretileceğini açıklar, quantum mekaniği bilimi atomaltı parçacıkların sahip oldukları özellikleri açıklar, termodinamik teorisi enerjinin ve mekanik işin sahip olduğu özellikleri açıklar. Üniversitede okutulan fizik bilimine ait giriş düzeyinde bir ders kitabı aşağı yukarı bu teorilerin her birine hasredilmiş ayrı bölümlerden mürekkeptir. Benzer teoriler, halihazırda, sosyal bilimlerde de mevcuttur. Mesela, psikoloji biliminde kognitif uyumsuzluk teorisi (*/cognitive dissonance theory*) insanların bir olaya ait gözlemlerinin o olaya ait beklentilerinden farklı

olduğu zaman nasıl reaksiyon gösterdiklerini açıklar, genel caydırıcılık teorisi (*/general dissonance theory*) niçin bazı insanların mesela yasadışı müzik indirme veya yazılım korsanlığı yapma gibi münasebetsiz veya kriminal davranışlara angaje olduklarını açıklar ve planlı davranış teorisi (*/theory of planned behavior*) insanların günlük yaşamlarında bilinçli akıl yürütmeli seçimleri nasıl yaptıklarını açıklar.

Bilimsel araştırmanın amacı yasaları ve postüla teorileri keşfetmektir ki bu teoriler doğal veya sosyal fenomenleri açıklayabilsin, ya da başka bir deyişle, bilimsel bilgiyi inşa edebilsin. Şunu anlamak önemlidir ki bu bilgi mükemmel olmayabilir veya hatta hakikatten (*/truth*) hayli uzak olabilir bile. Bazen, tek bir evrensel hakikat olmayabilir, daha ziyade “çoklu hakikatlerin” bir denkleşimi (*/equilibrium*) mevcut olabilir. Şunu anlamak zorundayız ki bilimsel bilginin üzerine oturduğu teoriler belirli bir fenomene dair bir bilimci tarafından teklif edilen açıklamalardır sadece. Böylece, iyi veya kötü açıklamalar olabilir, bu açıklamaların gerçeğe ne derece iyi uyduğuna bağlı olarak, ve bu da bizi iyi veya kötü teorilere götürür. Bilimin ilerlemesi demek, bizim daha isabetli enstrümanları kullanarak yaptığımız daha iyi gözlemler ve daha enforme edilmiş mantıksal akıl yürütme vasıtasıyla, zaman içerisinde daha kötü teorilerden daha iyi teorilere doğru ilerlememiz demektir.

Bilimsel yasalara ve teorilere bir mantık ve kanıt prosesinden geçerek ulaşırız. Mantık (teori) ve kanıt (gözlemler) iki, yalnızca iki, sütundur ki bilimsel bilgi onların üzerine oturmaktadır. Bilimde, teori ve gözlemler karşılıklı ilişkililerdir ve biri olmadan diğeri olamaz. Teoriler gözlediğimiz şeye mana (*/meaning*) ve anlam (*/significance*) kazandırır ve gözlemler mevcut teoriyi geçerlemeye (*/validate*) veya rafine etmeye ya da yeni bir teori inşa etmeye yardım eder. Herhangi bir başka vasıta ile bilgi edinmek, mesela inanç veya otorite gibi, bilim olarak telakki edilemez.

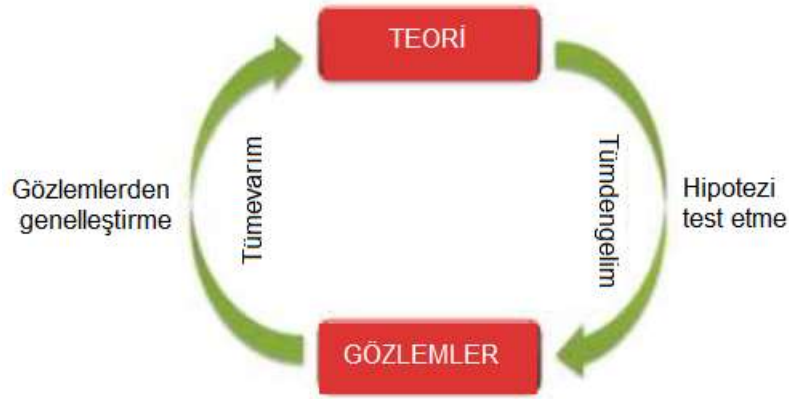
Bilimsel Araştırma

Teoriler ve gözlemler bilimin üzerlerine oturduğu iki sütun olduğuna göre, bilimsel araştırma iki düzeyde icra edilir: bir teorik düzey ve bir ampirik düzey. Teorik düzey doğal veya sosyal bir fenomen hakkında soyut kavramların geliştirilmesi ve bu kavramlar arasındaki ilişkiler (yani “teorilerin” inşası) ile alakalıyken, ampirik düzey teorik kavramların ve ilişkilerin, bizim gerçekteki gözlemlerimizi ne kadar iyi yansıttığını görmek için test edilmesi ile alakalıdır. Elbette nihai amaç daha iyi teoriler inşa etmektir. Zamanla, bir teori daha ve daha rafine olur (yani, gözlenen gerçekliğe daha iyi uyar) ve bilim olgunluk kazanır. Bilimsel araştırma teori ve gözlemler arasında sürekli geri ve ileri hareket etmeyi (yani iterasyonu) içerir. Hem teori ve hem de gözlemler bilimsel araştırmanın esansiyel bileşenleridir. Mesela, çıkarımda bulunmak için sırf gözlemlere dayanmak ve teoriyi gözardı etmek geçerli bilimsel araştırma olarak addedilmez.

Araştırmacının eğitime ve ilgisine bağlı olarak, bilimsel tahkikat iki olanaklı formdan birini alabilir: tümevarımcı veya tümdengelimci. **Tümevarımcı araştırmada**, araştırmacının amacı gözlenen veriden teorik kavramlar ve biçimler çıkarmaktır. **Tümdengelimci araştırmada**, araştırmacının amacı teoride bilinen kavram ve biçimleri yeni ampirik veri kullanarak test etmektir. Bu yüzden, tümevarımcı araştırma aynı zamanda *teori-inşa-edici* araştırma ve tümdengelimci araştırma *teori-test-edici* araştırma olarak adlandırılır. Dikkat edilmelidir ki teori-test-etmenin amacı yalnızca bir teoriyi test etmek değildir, fakat olanaklıca onu rafine etmek, geliştirmek ve genişletmektir. Figür 1.1 tümevarımcı ve tümdengelimci araştırmanın birbirini tamamlayıcı tabiatını remetmektedir. Dikkat edilmelidir ki tümevarımcı ve tümdengelimci araştırma, teori ve gözlemler arasında sürekli gidip-gelen (*/iteratif*) araştırma döngüsünün iki yarısıdır. Tümevarımcı veya tümdengelimci araştırma yapabilmeyiz, eğer

araştırmanın hem teori ve hem de veri bileşenlerine aşına değilsen. Tabiatıyla, tam bir araştırmacı o'dur ki bütün araştırma döngüsünü katedebilir ve hem tümevarımcı ve hem de tümden gelimci araştırmayı yürütebilir.

Şunu anlamak önemlidir ki hem teori-inşa-etme (tümevarımcı araştırma) ve hem de teori-test-etme (tümdengelimci araştırma) bilimin ilerlemesi için elzemdir. Seçkin teoriler değerli değildir, eğer gerçeklikle eşleşmiyorlar ise. Bunun gibi, veri dağları aynı zamanda işe yaramazdır, manalı teorilerin oluşturulmasına katkıda bulunabilene kadar. Figür 1.1 de görüldüğü üzere bu iki prosesi bir döngüsel ilişki içerisinde mülâhaza etmekten ziyade, belki de onları bir sarmal olarak mülâhaza etmek daha iyi olabilir, ki teori ve veri arasındaki her bir iterasyon ilgilenilen fenomenin daha iyi açıklamalarına ve daha iyi teorilere katkıda bulunmaktadır. Hem tümevarımcı ve hem de tümdengelimci araştırma her ne kadar bilimin ilerlemesi için önemli olsa da, şu açık ki öncül teoriler veya açıklamalar az sayıda olduğu zaman tümevarımcı (teori-inşa-edici) araştırma daha değerliken, aynı fenomen hakkında çok sayıda rakip teori olduğu ve araştırmacılar hangi teorinin hangi koşullar altında daha iyi işlediğini bilmekle ilgilendikleri zaman tümdengelimci (teori-test-edici) araştırma daha üretkendir.



Figür 1.1 Araştırma Döngüsü

Teori inşa etme ve teori test etme sosyal bilimlerde bilhassa zordur, şöyle ki teorik kavramların kesinliksiz mahiyeti, onları ölçecek araçların kifayetsizliği ve ilgilenilen fenomene aynı zamanda tesir edebilecek pek çok hesaba katılmayan faktörün mevcudiyeti bu zorluğun nedenleridir. Keza, işlemeyen teorileri çürütmek de çok zordur. Mesela, Karl Marx'ın komünizm teorisi ekonomik üretimin efektif bir vasıtası olarak on yıllarca ayakta kaldıktan sonra ancak, ekonomik gelişmenin ve sosyal refahın artırılmasında kapitalizm karşısında itibarını kaybetmiştir. Sovyetler Birliği ve Çin Halk Cumhuriyeti gibi bir zamanların komünist ekonomileri nihayetinde kar maksimizasyonuna dayalı özel sektör işletmeleri ile karakterize edilen daha kapitalist ekonomilere doğru yol aldılar. Mamafih, Amerika Birleşik Devletleri'nde ipotekli mülkiyet ve finans endüstrilerinin son zamanlardaki çöküşü, kapitalizmin de keza kendi eksiklikleri olduğunu ve ekonomik gelişmenin ve sosyal refahın beslenmesinde önceden zannedildiği kadar efektif olmadığını tanıtlamıştır. Doğa bilimlerindeki teorilere benzemezce, sosyal bilim teorileri nadiren mükemmeldir, ki bu da araştırmacılara o teorileri ıslah etmek ya da kendi yeni alternatif teorilerini inşa etmek için çok fazla fırsat verir.

Bilimsel araştırma yürütmek, dolayısıyla, teorik ve ampirik düzeylerde iş görmek için ihtiyaç duyulan, sırasıyla, iki grup beceri gerektirir – teorik ve metodolojik. Metodolojik beceriler (“know-how”) nispeten standarttır, disiplinler arasında değişiklik göstermez ve doktora programları vasıtasıyla kolayca kazanılabilir. Mamafih, teorik beceriler (“know-what”) ustalaşmakta hayli zordur, yıllar süren gözlem ve yansılama icabettirir ve öğretilen bilmekten daha ziyade deneyimle öğrenilebilen zımni becerilerdir. İnsanoğlunun tarihindeki en büyük

bilimcilerin tümü, mesela Galileo, Newton, Einstein, Neils Bohr, Adam Smith, Charles Darwin, ve Herbert Simon gibi, usta teorisyenlerdi ve postüla ettikleri teorilerle anımsanırlar ki o teoriler bilimin gidişatını değiştirmiştir. Metodolojik becerilere sıradan bir araştırmacı olmak için ihtiyaç duyulur, fakat sıradışı bir araştırmacı olmak için teorik becerilere ihtiyaç vardır!

Bilimsel Metot

Önceki kısımlarda, bilimi bilimsel metot vasıtasıyla edinilen bilgi olarak betimledik. Peki “bilimsel metot” tam olarak nedir? **Bilimsel metot** bilimsel bilginin inşasına yarayan standartlaştırılmış bir teknikler kümesine tekabül eder, mesela geçerli gözlemlerin nasıl yapılacağı, sonuçların interpretasyonunun nasıl yapılacağı ve o sonuçların nasıl genelleştirileceği gibi. Bilimsel metot araştırmacılara önceki teorileri ve önceki bulguları bağımsızca ve tarafsızca test edebilme ve onları açıkça tartışabilme, modifiye edebilme ve iyileştirebilme olanağı sağlar. Bilimsel metot esasen dört karakteristiğe sahiptir:

- *Mantıklı*: Bilimsel çıkarımlar akıl yürütmenin mantıksal prensiplerine dayanmalıdır.
- *Onaylanabilir*: Türetilen çıkarımlar gözlenen kanıt ile eşleşmelidir
- *Tekrarlanabilir*: Başka bilimciler bir bilimsel öğrenme-çalışmasını bağımsızca yinelemeli veya tekrarlamalı ve aynı değilse bile benzer sonuçlara ulaşmalıdır.
- *Eleştirilebilir*: Kullanılan prosedürler ve türetilen çıkarımlar başka bilimciler tarafından yapılacak eleştirel incelemeye (peer review) karşı durabilmelidir.

Herhangi bir tahkikat alanında, eğer o alanın temel yasalarının veya teorilerinin bilimsel metoda (/scientific method) uygun olarak test edilmesine olanak sağlanmıyor ise, o tahkikat alanı “bilim (/science)” diye adlandırılabilir. Mesela, teoloji (dine ait öğrenme-çalışması) bilim değildir çünkü teolojik fikirler (mesela Tanrı’nın varlığı gibi) mantıklı, onaylanabilir, tekrarlanabilir ve eleştirilebilir bir metot kullanılarak bağımsız gözlemcilerce test edilemez. Benzerce, sanat, müzik, edebiyat, felsefe, tarih, hukuk, İngiliz dili, Türk dili vesaire (/arts and humanities) keza bilim olarak addedilmez, her ne kadar onlar da kendilerince yaratıcı ve değerli uğraşlar olsalar dahi.

Bilimsel metot, sosyal bilimlerde uygulandığı üzere, nitel ve nicel veri toplamak ve analiz etmek için çeşitli araştırma yaklaşımlarını, araçları ve teknikleri ihtiva eder. Bunlar laboratuvar deneyleri, saha yoklamaları, vaka araştırması, etnoğrafik araştırma, eylem araştırması vesairedir. Bu kitap çoğunlukla bu farklı metotları öğrenmeye adanmıştır. Mamafih, şunu iyi anlamak lazım ki bilimsel metot araştırmanın başlıca ampirik düzeyinde iş görür, yani, gözlemlerin nasıl yapılacağına ve yapılan gözlemlerin nasıl analiz edileceğine. Araştırmanın teorik düzeyi ile, bilimsel metodun doğrudan çok az münasebeti vardır, ki orası bilimsel araştırmanın gerçekten en meydan okuyucu kısmıdır.

Bilimsel Araştırma Tipleri

Araştırmanın maksadına bağlı olarak, bilimsel araştırma projeleri üç tipte gruplandırılabilir: keşifleyici (/exploratory), betimleyici (/descriptive) ve açıklayıcı (/explanatory). **Keşifleyici araştırma** sıklıkla yeni tahkikat alanlarında yürütülür, ki araştırmacı şunları amaçlar: (1) belirli bir fenomenin, problemin veya davranışın cesametini ve şümülünü tespit etmek, (2) o fenomen hakkında bazı başlangıçsal fikirler (veya “sezgiler”) üretmek ya da (3) o fenomene dair daha yoğun bir öğrenme-çalışmasına girişmenin fizibilitesini test etmek. Mesela, bir ülkenin halkı, bir ekonomik durgunluk esnasında, hükümet politikalarından genellikle memnun değiller ise, keşifleyici araştırma halkın memnuniyetsizlik derecesini ölçmeye, böylesi bir memnuniyetsizliğin nasıl dışa vurulduğunu (mesela toplu protesto gösterilerinin frekansı gibi)

ve böylesi bir memnuniyetsizliğin farazi nedenlerini (mesela enflasyonu, faiz oranlarını, işsizliği veya yüksek vergileri kontrol altında tutmaya yönelik etkisiz hükümet politikaları gibi) anlamaya yönlilebilir. Böyle bir araştırma, üçüncü-parti kaynaklarca arşivlenmiş olan kamusal raporlarda yer alan mesela ekonomik indikatörlerin kestirimleri (gayrisafi yurtiçi hasıla, işsizlik, tüketici fiyat endeksi, vesaire) gibi rakamların incelenmesini ihtiva edebilir, ki bu rakamlar konunun uzmanlarıyla, üst düzey ekonomistlerle veya yetkili hükümet üyeleriyle yapılan mülakatlardan elde edilir ve/veya benzer problemlerin ele alınmış olduğu geçmişteki misaller üzerinde yapılan öğrenme-çalışmalarından türetilir. Bu araştırma hedef problemin çok isabetli olarak anlaşılmasını temin edemeyebilir, fakat problemin mahiyetinin ve şümülünün tespitine yarar ve daha derinlemesine araştırma için yararlı bir prekürsör olur.

Betimleyici araştırma ilgilenilen bir fenomenin dikkatli gözlemlerini ve detaylı dökümantasyonunu yapmaya yöneliktir. Bu gözlemler bilimsel metoda dayanmalıdır (yani, tekrarlanabilir, kesin, vesaire olmalıdır) ve bu yüzden, eğitimsiz kimselerin yapmış olduğu öylesine gözlemlerden daha güvenilirlerdir. Betimleyici araştırmanın misalleri Türkiye İstatistik Kurumu tarafından demografik istatistiklerin veya Türkiye İş Kurumu tarafından istihdam istatistiklerinin tablolanmasıdır, ki bu kurumlar nüfus sayımları veya istihdam yoklamaları üzerinden cinsiyete göre nüfus artışı veya sektörlere göre istihdamı kestirmek için aynı veya benzer enstrümanları kullanırlar. Eğer ölçme enstrümanlarında herhangi bir değişiklik yapılırsa, değiştirilmiş ve değiştirilmemiş estrümantasyona ait kestirimler beraber sunulur ki okuyucu nüfus veya istihdam trendlerinin önce-ve-sonra karşılaştırmasını doğru bir şekilde yapabilsin. Başka betimleyici araştırmalara misal olarak kentsel nüfus içerisinde ergen gençler arasındaki çete faaliyetlerinin tarih-sıralı etnoğrafik raporları; seçilmiş topluluklarda dinsel, kültürel veya etnik pratiklerin süregenliği veya evrimi; ve Afrika ülkelerinde demokrasi hareketlerinin yayılmasında mesela Twitter ve anlık mesajlaşma gibi teknolojilerin rolü verilebilir.

Açıklayıcı araştırma gözlenen fenomenlerin, problemlerin veya davranışların açıklamalarını bulmaya çalışır. Betimleyici araştırma bir fenomenin ne olduğunu, nerede olduğunu ve ne zaman olduğunu incelerken, açıklayıcı araştırma niçin ve nasıl tipindeki soruların cevaplarını bulmaya çalışır. Hedef fenomenin nedensel faktörlerini ve çıkımlarını (/outcomes) teşhis etmek suretiyle, araştırmada “noktaları birleştirmeye” uğraşır. Mesela, ergenlikte suç işlemenin veya çete şiddetinin arkasındaki sebepleri anlamak, böylesi toplumsal rahatsızlıkların üstesinden gelmek için stratejiler reçete etmek amacıyla yürütülebilecek açıklayıcı tipte bir araştırma olurdu. Akademi veya doktora düzeyindeki araştırmaların çoğu açıklayıcı tiptedir, her ne kadar akademik araştırmanın başlangıç safhalarında bir miktar keşifleyici ve/veya betimleyici araştırmaya ihtiyaç duyulabilse bile. Gözlenen olaylar için açıklamalar bulmaya çalışmak sezgi, nüfuz-u nazar ve kişisel deneyimin yanısıra güçlü teori ve interpretasyon becerileri gerektirir. Bunları iyi yapabilenler aynı zamanda kendi disiplinlerinde dünyanın en çok takdir edilen bilimcileridir.

Bilimsel Düşüncenin Tarihi

Bu bölümü bitirmeden önce, tarihte geriye gitmek ve bilimin zaman içerisinde nasıl evrilmiş olduğunu görmek ve bu evrime katkıda bulunmuş olan kilit bilimsel zihinleri tanımak ilginç olabilir. Her ne kadar bilimsel ilerlemenin adımları pek çok yüzyıldır kaydedilmekte olsa bile, “bilim (/science)”, “bilimci (/scientist)” ve “bilimsel metot (/scientific method)” terimleri ancak 19 uncu yüzyılda ortaya çıkmıştır. O zamana kadar, bilim felsefenin bir branşı olarak görüldü ve felsefenin mesela mantık, metafizik, etik ve estetik gibi branşlarıyla beraber varlığını sürdürdü; tabii ki bu branşların bazıları arasındaki sınırların bulanık olduğunu da biliyoruz.

İnsanın tahkikata başladığı ilk zamanlarda, bilgi genellikle inanca dayalı teolojik buyruklar cinsinden mülahaza edildi. Buna ilk meydan okuyan M.Ö. 3 üncü yüzyılda mesela Socrates, Plato ve Aristotle gibi Yunan filozofları oldu ki onlar varlığın ve dünyanın temel mahiyetinin **rasyonalizm** adı verilen sistematik mantıksal akıl yürütme prosesi ile daha isabetlice anlaşılabilceğini ileri sürdüler. Bilhassa, Aristotle’ın klasik çalışması *Metaphysics* (kelimesi kelimesine “fiziksel [varlığın] ötesi” demek) *teolojiyi* (Tanrıları öğrenme-çalışması) *ontolojiden* (oluşu (/being) ve varlığı (/existence) öğrenme-çalışması) ve *universal bilimden* (mantığın üzerine oturduğu ilk prensipleri öğrenme-çalışması) ayırdı. Rasyonalizm (“rasyonalite” ile karıştırılmamalı) akıl yürütmeyi bilginin veya gerekçelendirmenin kaynağı olarak kabul eder ve şunu ileri sürer ki hakikatin kriteri duysal değil fakat daha ziyade akılsal ve tümdengelimcidir, öyle ki bu kriter çoğu zaman bir ilk prensipler veya aksiyomlar kümesinden türetilir (mesela Aristotle’ın çelişkisizlik yasası” gibi).

Bilimsel düşüncede ikinci büyük dönüşüm 16 ncı yüzyılda meydana geldi, öyle ki o zaman İngiliz filozof Francis Bacon (1561-1626) bilginin ancak gerçek dünyada yapılan gözlemlerden türetilbileceğini ileri sürdü. Bu öncüle dayanarak, Bacon bilgi edinmenin (bir akıl yürütme faaliyeti olmaktan daha ziyade) bir ampirik faaliyet olduğunu vurguladı ve felsefenin tesirli bir branşı olarak **ampirizmi** geliştirdi. Bacon’un çalışmaları bilimsel tahkikatın tümevarımcı metodlarının popüler olmasına ve sistematik gözlemi, ölçmeyi ve deneylemeyi ihtiva eden “bilimsel metodun” (orjinalinde “Baconian method” denir) gelişmesine yol açtı. Öte taraftan, hatta, belki de, artık insanların günlük yaşantısına musallat olan kilise buyruklarının veya teolojik akidelerin “gözlenebilir” oldukları için reddine kapı aralamış oldu.

Ampirizmin rasyonalizmle çatışması Orta Çağ boyunca sürdü, ki bu zaman zarfında filozoflar geçerli bilgi ediminin en efektif yolunu bulmaya çalışıyorlardı. Fransız filozof Rene Decartes (1596-1650) rasyonalistlerin yanında yer alırken İngiliz filozoflar John Locke (1632-1704) ve David Hume (1711-1776) ampiristlerin yanında yer aldılar. Mesela Galileo Galilei (1564-1642) ve Sir Issac Newton (1643-1727) gibi bilimciler spesifik olarak doğayı ve fiziksel evreni anlamaya odaklanmak için bu iki fikri kaynaştırarak bir **doğa felsefesi** yaratmaya giriştiler, ki doğa bilimlerinin prekürsörü olarak telakki edilir. Galileo doğanın yasalarının matematiksel olduğunu ifade eden belki de ilk kişidir ve deneylemenin ve matematiğin innovatif bir kombinasyonu vasıtasıyla astronomi alanına katkıda bulunmuştur.

On sekizinci yüzyılda, Alman filozof Immanuel Kant (1772-1804), *Critique of Pure Reason* adlı kitabında, deneyimin salt subjektif olduğunu ve deneyimlerin subjektif doğasını öncelikle derinlemesine incelemeyi, onları saf akıl yürütme vasıtasıyla işlemenin teorik illüzyonlara yol açacağı argümanını savunarak, ampirizm ve rasyonalizm arasındaki anlaşmazlığa bir çözüm bulmaya çalıştı. Kant’ın fikirleri **Alman idealizminin** gelişiminin yolunu açtı, ki o da mesela fenomenoloji, hermenötik ve eleştirel sosyal teori gibi interpretif tekniklerin gelişmesine tesir etti.

Aynı zaman zarfında, Fransız filozof Auguste Comte (1798-1857), sosyoloji disiplininin kurucusu olarak, rasyonalizmi ve ampirizmi **pozitivizm** adı verilen yeni bir doktrin içerisinde katıştırmaya girişti. Öne sürdüğü şeydi ki teoremin ve gözlemlerin birbirlerine sirküler bağımlılığı vardır. Akıl yürütme vasıtasıyla yaratılabilen teoriler, ancak gözlemler vasıtasıyla doğrulanabilirlerse otantiklerdir. Doğrulamanın önemini kabul edilmesi modern bilimlerin felsefeden ve metafizikten ayrılması sürecini başlattı ve daha sonra “bilimsel metodun” bilimsel iddiaları geçerlemenin başlıca vasıtası olarak yolunu açtı. Comte’un fikirleri Emile Durkheim (1858-1917) tarafından sosyolojik pozitivizmin (sosyal araştırmanın üzerine oturtulacağı bir zemin olarak pozitivizm) geliştirilmesinde ve Ludwig Wittgenstein (1889-1951) tarafından mantıksal pozitivizmin geliştirilmesinde daha ileriye taşındı.

Yirminci yüzyılın başlarında, pozitivizmin güçlü çıkışları düşüncede Alman idealizmi ekolüne

bağlı interpretif sosyologlar tarafından (antipozitivistler) reddedildi. **Antipozitivizm** mesela yapılandırılmamış görüşmeler ve katılımcı gözlem gibi nitel metotları istihdam ederken, pozitivizm tipik olarak mesela deneyler ve yoklamalar (/surveys) gibi nicel araştırma metotları ile ve bariz felsefi yükümlülüklerden muafiyet ile eşleştirildi. Hatta mesela Amerikan sosyolojist Paul Lazarsfield (1901-1976) gibi pozitivizmin pratisyenleri dahi, gözlemci yanlılığını ve pozitivist tahkikatta yapısal kısıtlamaları potansiyel problemler olarak kabul etmişlerdir, ki Lazarsfield büyük-ölçekli yoklama araştırmasının ve yoklama verisinin analizi için istatistiksel tekniklerin öncülüğünü yapmıştır. Karşı tarafta, antipozitivistler şunu vurguladılar ki sosyal eylemler bireylerin kendi kişisel eylemlerine atfettikleri mana ve maksatların bir anlayışına dayalı interpretif araçlar vasıtasıyla öğrenilmeye-çalışılmalıdır, ki bu fikir Georg Simmel'in (1858-1918) sembolik interaksyonizm üzerindeki çalışmasına, Max Weber'in (1864-1920) ideal tipler üzerindeki çalışmasına ve Edmund Husserl'in (1859-1938) fenomenoloji üzerindeki çalışmasına esin kaynağı olmuştur.

Yirminci yüzyılın ortalarından sonlarına doğru, hem pozitivist ve hem de antipozitivist düşünce ekolleri eleştirilere ve modifikasyonlara maruz kaldılar. İngiliz filozof Karl Popper (1902-1994) insan bilgisinin meydan okunamaz, çok katı temeller üzerine oturmaktan daha ziyade asla katiyetle ispatlanılabilecek fakat ancak yanlışlanabilecek tentatif konjektürlerin bir kümesi üzerine oturduğunu öne sürdü. Ampirik kanıt bu konjektürleri veya "teorileri" yanlışlamanın mesnedidir. **Postpozitivizm** (ya da postampirizm) olarak adlandırılan bu metateorik duruş, şunu öne sürmek suretiyle pozitivizmi ıslah eder ki her ne kadar yanlış inançları reddetmek olanaklı olsa dahi, hakikati doğrulamak olanaksızdır; ve aynı zamanda pozitivist bir objektif hakikat nosyonuna sahip çıkmayı ve onun bilimsel metot üzerindeki önemini vurgulamayı da sürdürür.

Benzerce, antipozitivistler de keza eleştirilmişlerdir, şöyle ki yalnızca toplumu anlamaya çalışmakla fakat toplumu eleştirmemekle ve daha iyi yönde değiştirmemekle. Bu düşüncenin kökleri Alman filozoflar Karl Marx (1818-1883) ve Fredrich Engels (1820-1895) tarafından yazılan *Das Capital* adlı kitaptan neşet eder ki kapitalist toplumları sosyal olarak eşitsiz ve etkinsiz olmakla eleştirmiş ve bu eşitsizliğin sınıf çatışması ve proleter devrimler vasıtasıyla çözümlenmesini önermiştir. Marksizm mesela Almanya, İtalya, Rusya ve Çin gibi ülkelerdeki sosyal dönüşümlere/devrimlere esin kaynağı oldu, fakat arzulanan sosyal eşitliğe ulaşmakta genellikle başarısız oldu. Yirminci yüzyılda, Max Horkheimer (1895-1973) ve Jürgen Habermas (d. 1929) tarafından ileri sürülen **eleştirel araştırma** (aynı zamanda eleştirel teori diye adlandırılır) sosyal eşitsizliği eleştirmeye ve çözmeye dair benzer fikirleri devam ettirmekle beraber şunu da ilave eder ki halk (/insanlar/people) kendi sosyal ve ekonomik şartlarını değiştirmek için bilinçlice davranabilmelidir ve davranmalıdır, ki her ne kadar öyle yapabilirlikleri sosyal, kültürel ve politik baskınlığın çeşitli formlarınca kısıtlanmış olsa dahi. Eleştirel araştırma çağdaş toplumdaki karşıtlıkları, çatışmaları ve çelişkileri analiz etmek suretiyle, statükonun sınırlayıcı ve yabancılaştırıcı koşullarını ortaya çıkarmaya ve eleştirmeye girişir ve yabancılaştırmanın ve baskınlığın nedenlerini elemine etmeye (yani, ezilen sınıfı kurtarmaya) çabalar. Bu farklı araştırma felsefelerinin ve yaklaşımlarının daha fazlası bu kitabın sonraki bölümlerinde kapsanacaktır.

Bölüm 2

Bir Araştırmacı Gibi Düşünmek

İyi araştırma yürütmek, kişinin önce beynini bir araştırmacı gibi düşünecek tarzda yeniden eğitmesini gerektirir. Bu da fiili gözlemlerin bir soyutlamasını zihinde canlandırmayı, saklı kavramları (/concepts) ve biçimleri (/patterns) teşhis etmek için zihninde “noktaları birleştirmeyi” ve o biçimlerden ilk gözlem sahasının ötesinde başka bağlamlara uygulanacak yasalar ve teoriler sentezlemeyi gerektirir. Araştırma gözlemlerin yapıldığı ampirik bir düzlem ve o gözlemlerin genelleştirilebilir yasalara ve teorilere soyutlandığı bir teorik düzlem arasında sürekli geri ve ileri hareket etmeyi ihtiva eder. Bu öyle bir yetenektir ki geliştirilmesi yıllar alır ve yüksek lisans veya doktora programlarındaki derslerde öğretilen ya da endüstri eğitimleriyle kazanılabilen bir şey değildir; ve doktora öğrencilerinin en büyük açığı da budur. Bir araştırmacı gibi düşünmek için gereksinim duyulan zihinsel soyutlamaların bazıları analiz birimi, yapılar, hipotezler, operasyonelleştirme, teoriler, modeller, tümevarım, tümdengelim vesairesidir ki onları bu bölümde inceleyeceğiz.

Analiz Birimi

Herhangi bir sosyal bilim araştırmasında verilecek ilk kararlardan biri bir bilimsel öğrenme-çalışmasının birimidir. **Analiz birimi** soruşturmanın hedefi olan kişi, topluluk veya objedir. Tipik analiz birimleri bireyleri, grupları, organizasyonları, ülkeleri, teknolojileri, objeleri vesaire ihtiva eder. Mesela, eğer insanların alışveriş davranışını, öğrenme çıkımlarını (/outcome) veya yeni teknolojilere karşı tutumlarını öğrenmeye-çalışmakla ilgileniyorsak, o zaman analiz birimi *bireydir*. Eğer sokak çetelerinin veya organizasyonlarda çalışan takımların karakteristiklerini öğrenmeye-çalışmak istiyorsak, o zaman analiz birimi *gruptur*. Eğer araştırmanın amacı firmaların karlılıklarını nasıl artıracaklarını veya iyi yönetimsel kararlar verebileceklerini anlamak ise, o zaman analiz birimi *firmedir*. Bu durumda, kararlar o firmalardaki bireyler tarafından veriliyor olsa dahi, o bireylerin kendi kararlarından ziyade mensubu oldukları firmanın kararı olarak kabul edilir. Eğer araştırma ulusal kültürlerin farklılıklarını anlamaya yönelikse, o zaman analiz birimi *ülkedir*. Hatta cansız nesneler bile analiz birimleri olarak hizmet edebilir. Mesela, eğer bir araştırmacı web sayfalarının kullanıcılar için nasıl daha çekici hale geleceğini anlamakla ilgileniyorsa, o zaman analiz birimi *web sayfasıdır* (kullanıcılar değil). Eğer firmalar arasında karşılıklı bilgi transferinin nasıl olduğunu öğrenmeye-çalışmayı arzu ediyorsak, o zaman analiz birimiz *firma çifti* (bilgi gönderen ve alan firmaların kombinasyonu) olur.

Analiz birimini anlamak bazen hayli kompleks olabilir. Mesela, eğer belli mahallelerde suç işleme oranlarının niçin daha yüksek olduğunu öğrenmeye-çalışmayı arzu ediyorsak, o zaman analiz birimiz *mahalledir*, ne işlenen suçlar ve ne de o suçları işleyenler değildir. Çünkü bizim tahkikatımızın objesi suçlular değil mahallelerdir. Mamafih, eğer farklı mahallelerde mesela cinayet, soygun, tecavüz vesaire gibi farklı tipte suçları karşılaştırmak istersek, o zaman analiz birimiz *suç* olur. Eğer suçluların niçin yasadışı faaliyetlere bulaştıklarını öğrenmeye-çalışmayı arzu ediyorsak, o zaman analiz birimi *birey* (yani, suçlu) olur. Benzerce, eğer niçin bazı inovasyonların diğerlerine göre daha başarılı olduklarını öğrenmeye-çalışmak istiyorsak, analiz birimiz bir *inovasyondur*. Mamafih, eğer bazı organizasyonların diğerlerine göre nasıl daha istikrarlıca inovasyon yaptıklarını öğrenmeye-çalışmak arzusundaysak, o zaman analiz birimi *organizasyondur*. Böylece, aynı araştırma öğrenme-çalışması içerisinde yer alan birbiriyle ilişkili iki araştırma sorusu tamamen birbirinden farklı analiz birimlerine sahip

olabilirler.

Analiz birimini anlamak önemlidir, çünkü öğrenme-çalışman için ne çeşit veri toplayacağını ve onu kimden toplayacağını şekillendirir. Eğer analiz birimin bir web sayfası ise, mevcut veri sayfalarından web sayfaları hakkında veri toplarsın, web sayfalarını nasıl kullandıkları hakkında insanları yoklamazsın. Eğer analiz birimin organizasyon ise, o zaman mesela organizasyonun hacmi, gelirleri, hiyerarşisi veya absorpsiyon kapasitesi gibi organizasyon düzeyinde değişkenleri ölçüyor olursun. Bu veri mesela finansal kayıtlar veya (kendinden ziyade) organizasyonu temsil ettiği kabul edilen tepe yöneticilerle (CEO) yapılan yoklamalar gibi çeşitli kaynaklardan gelebilir. Mesela CEO maaşı gibi bazı değişkenler birey düzeyindeymiş gibi görünebilir, fakat olguda/gerçekte, organizasyon düzeyinde bir değişken olabilir aynı zamanda, çünkü her organizasyonda aynı anda yalnızca bir CEO maaşı ödenir. Bazen, daha düşük bir analiz birimi düzeyinde veri toplandıktan sonra, o veriyi daha yüksek bir analiz birimi düzeyinde toplatırmak olanaklıdır. Mesela, eğer organizasyonlarda takım çalışmasını öğrenmeye-çalışıyorsak, farklı organizasyon takımlarının üyelerini bireysel olarak yoklayabilir ve mesela uyum ve çatışma gibi takım düzeyindeki değişkenlere ait takım düzeyinde kompozit bir skor elde etmek için takım üyelerinin bireysel skorlarının ortalamasını alabiliriz. Sonraki kısımda “değişken” kavramını derinlemesine inceleyeceğiz.

Kavramlar, Yapılar ve Değişkenler

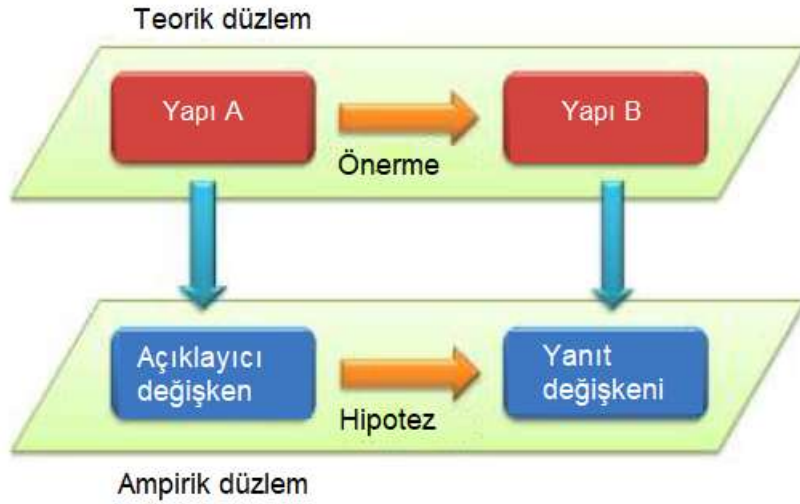
Bölüm 1 de tartıştığımız üzere araştırma keşifleyici, betimleyici veya açıklayıcı olabilse dahi, çoğu bilimsel araştırma açıklayıcı tiptedir öyle ki gözlenen doğal veya sosyal fenomenlerin potansiyel açıklamalarını bulmaya çalışır. Açıklamalar nesnelere, olaylara veya insanlara dair **kavramların** (/concepts) veya genelleştirilebilir özelliklerin veya karakteristiklerin geliştirilmesini icap ettirir. Mesela bir kişi, bir firma, veya bir otomobil gibi nesneler birer kavram değildir, fakat mesela bir kişinin mültecilere karşı tutumu, bir firmanın inovasyon kapasitesi ve bir otomobilin ağırlığı gibi onların spesifik karakteristikleri veya davranışları birer kavram olarak görülebilir.

Bilerek veya bilmeyerek, gündelik konuşmalarımızda farklı çeşitlerde kavramlar kullanırız. Bu kavramların bazıları müşterek dilimiz vasıtasıyla zaman içerisinde geliştirilmiştir. Bazen, ilgilenilen bir fenomeni açıklamak için başka disiplinlerden veya dillerden kavramlar ödünç alırız. Mesela, fizikten ödünç alınan *çekim* fikri ticarete insanların niçin tercih ettikleri alışveriş yerlerine/destinasyonlarına “çekilme” eğilimi gösterdiklerini betimlemek için kullanılabilir. Onun gibi, *mesafe* kavramı iki farklı eşyerleşimli birey arasındaki sosyal ayrımın derecesini açıklamak için kullanılabilir. Bazen, önceki araştırmalarda henüz betimlenmemiş olan bir karakteristiği betimlemek için kendi kavramlarımızı yaratırız. Mesela, *teknostres* yeni bir kavramdır ki bir kişiden yeni bir teknolojiyi öğrenmesi istendiği zaman, o kişinin maruz kalabileceği zihinsel stresi ifade eder.

Kavramlar aynı zamanda soyutlamanın ileri düzeylerine geçebilirler. Mesela bir insanın *ağırlığı* gibi kavramlar kesin ve nesnelken, mesela bir insanın *kişiliği* gibi başka kavramlar çok daha soyut ve zihinde canlandırması çok daha zor olabilmektedir. Bir **yapı** (/construct) soyut bir kavramdır öyle ki belli bir fenomeni açıklamak için spesifik olarak seçilir (veya “yaratılır”). Bir yapı basit bir kavram olabilir, mesela bir kişinin *ağırlığı* gibi; veya birbiriyle ilişkili bir kavramlar kümesinin bir kombinasyonu olabilir, mesela bir kişinin *iletişim yeteneği* gibi, ki mesela kişinin sözcüklüğü, sözcüzimi ve harfdizimi gibi alttayatan çeşitli kavramlardan müteşekkildir. Misallerin ilki (*ağırlık*) bir **tek-boyutlu yapı** iken, sonraki (*iletişim yeteneği*) bir **çok-boyutlu yapıdır** (yani, çoklu alttayatan kavramlardan müteşekkildir). Yapılar ve kavramlar arasındaki ayrım çok-boyutlu yapılarda daha barizdir ki daha yüksek dereceli soyutlama bir

yapı olarak adlandırılır ve daha düşük dereceli soyutlamalar birer kavram olarak adlandırılırlar. Mamafih, bu ayrım tek-boyutlu yapılarda muğlaklaşır.

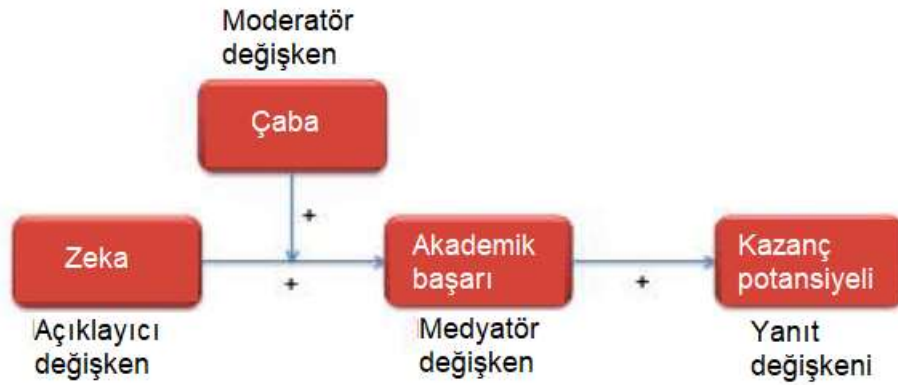
Bilimsel araştırma için kullanılan yapılar öyle kesin ve açık tanımlar olmalıdır ki onları kullanabilecek başkaları tarafından tam olarak ne manaya geldikleri ve ne manaya gelmedikleri anlaşılabilir. Mesela *gelir* gibi görünürde basit bir yapı aylık veya yıllık geliri, vergi-öncesi veya vergi-sonrası geliri ve kişi-başına veya aile-içi geliri ifade edebilir ve bu yüzden ne kesindir ne de açıktır. Tanımların iki tipi vardır: sözlük tanımları ve operasyonel tanımlar. Daha çok aşına olduğumuz sözlük tanımında, bir yapı sıklıkla bir sinonimi cinsinden tanımlanır. Mesela *tutum* bir duruş, bir davranış veya bir tavır olarak tanımlanır; ve diğer taraftan *tavır* bir tutum olarak tanımlanır. Döngüsel bir mahiyetteki bu tanımlar bilimsel araştırmada bir yapının manasını ve içeriğini işleyebilmek için bilhassa yararlı değildir. Bilimsel araştırma **operasyonel tanımları** gerektirir ki onlar yapıları ampirik olarak nasıl ölçülecekleri cinsinden tanımlar. Mesela *sıcaklık* gibi bir yapının operasyonel tanımı sıcaklığı Celcius ile mi, Fahrenheit ile mi yoksa Kelvin ile mi ölçmeyi planladığımızı spesifik olarak belirtmelidir. Mesela *gelir* gibi bir yapı aylık mı yoksa yıllık gelire mi, vergi-öncesi mi yoksa vergi-sonrası gelire mi ve kişi-başına mı yoksa aile-içi gelire mi ilgilendiğimiz cinsinden tanımlanmalıdır. Şimdi, mesela öğrenme, kişilik ve zeka gibi yapıları operasyonel olarak tanımlamanın ne kadar zor olabileceğini hayal et.



Figür 2.1 Araştırmanın teorik ve ampirik düzlemleri

Bir yapı ile sıkı bağlantılı ve bazen birbirini yerine kullanılan bir terim değişkendir (/variable). Etimolojik olarak, bir değişken öyle bir niceliktir ki değişebilir (yani, alçaktan yükseğe, negatiften pozitive vesaire); aksine, sabiteler (/constants) değişmezler (yani, sabit kalırlar). Mamafih, bilimsel araştırmada, bir **değişken** soyut bir yapının ölçülebilir bir temsilidir. Soyut varlıklar olarak, yapılar doğrudan ölçülmez, ve bu yüzden, değişken adı verilen vekil ölçüler bulmaya çalışırız. Mesela, bir kişinin *zekası* çoğu zaman onun *IQ* (/intelligence quotient) *skoru* ile ölçülür ki insanların tabii tutulduğu analitik ve biçim-eşlemeli bir testten üretilen bir endekstir. Bu durumda, *zeka* bir yapıdır ve *IQ skoru* zeka yapısını ölçen bir değişkendir. IQ skorlarının bir kişinin zekasını hakikaten ölçüp ölçmediği, pek çok kişi ölçtüğüne inansa dahi, katiyetle söylenemez; ve zekayı ne kadar iyi ölçtüğüne bağlı olarak, IQ skoru zeka yapısının iyi veya kötü bir ölçüsü olabilir. Figür 2.1 de gösterildiği üzere, bilimsel araştırma iki düzlem boyunca ilerler: bir teorik düzlem ve bir ampirik düzlem. Yapılar teorik (soyut) düzlemde kavramsallaştırılırken, değişkenler ampirik (gözlemsel) düzlemde operasyonelleştirilir ve ölçülür. Bir araştırmacı gibi düşünmek bu iki düzlem arasında geri ve ileri hareket edebilmek demektir.

Hangi niyetle kullanıldıklarına bağlı olarak, değişkenler açıklayıcı değişkenler (ya da bağımsız değişkenler), yanıt değişkenleri (ya da bağımlı değişkenler), medyatör değişkenler, moderatör değişkenler, veya kontrol değişkenleri olarak sınıflandırılabilir. Başka değişkenleri açıklayan değişkenlere **açıklayıcı değişkenler** (bağımsız değişken) denir; başka değişkenler tarafından açıklanan değişkenlere **yanıt değişkenleri** (bağımlı değişken) denir; yanıt değişkenlerini açıklıyorken, aynı zamanda açıklayıcı değişkenler tarafından açıklanan değişkenlere **medyatör değişkenler** (ya da vasati/aracı değişkenler) denir; ve açıklayıcı değişkenler ve yanıt değişkenleri arasındaki ilişkiye tesir eden değişkenlere **moderatör değişkenler** denir. Bir misal vermek gerekirse, eğer yüksek zekanın öğrencilerde daha iyi öğrenmeye neden olduğunu ifade edersek, o zaman zeka bir açıklayıcı değişken ve öğrenme bir yanıt değişkenidir. Bir yanıt değişkenini açıklamakta kullanmak istemediğimiz, fakat o yanıt değişkenine az çok etki edebilen başka ekstra değişkenler olabilir, ki neredeyse her zaman ziyadesiyle mevcuttur. Bir bilimsel öğrenme-çalışmasında, açıklayıcı değişkenin yanıt değişkenine yaptığı etkiyi ortaya çıkarabilmek için, bu ekstra değişkenlerin kontrol altına alınmaları gerekir.



Figür 2.2 Yapıların bir nomolojik ağı

Bu farklı değişken tipleri arasındaki farkları anlamak için, Figür 2.2 de gösterilen misali mülâhaza edelim. Eğer zekanın öğrencilerin akademik başarısına tesir ettiğine (ya da açıkladığına) inanıyorsak, o zaman mesela *IQ skoru* gibi bir zeka ölçüsü açıklayıcı değişkenimiz ve mesela *not ortalaması* gibi bir akademik başarı ölçüsü yanıt değişkenimizdir. Eğer akademik başarı üzerine zekanın etkisinin aynı zamanda öğrencinin öğrenme prosesine harcadığı çabaya bağlı olduğuna (yani, eşit zekaya sahip iki öğrenciden, daha çok çaba harcayanının daha az çaba harcayanına göre daha yüksek akademik başarıya ulaşacağına) inanıyorsak, o zaman *çaba* bir moderatör değişken olur. Bir başkası, öte taraftan, çabayı bir açıklayıcı değişken ve zekayı bir moderatör değişken olarak görebilir. Eğer akademik başarı daha yüksek kazanç potansiyeline doğru bir ara basamak olarak görülüyorsa, o zaman *akademik başarı* açıklayıcı değişkeni için yanıt değişkeni *kazanç potansiyeli* olur; ve zeka ve kazanç potansiyeli arasındaki ilişkide akademik başarı medyatör değişken olur. Böylece, değişkenlerin açıklayıcı değişken, yanıt değişkeni, moderatör değişken veya medyatör değişken olarak tanımlanması birbirleriyle olan ilişkilerinin mahiyetine bağlıdır. Birbirleriyle bağlantılı yapıların bir kümesi arasındaki ilişkileri tümüyle kapsayan bir ağı **nomolojik ağ** adı verilir (Figür 2.2). Bir araştırmacı gibi düşünmek yalnızca gözlemleri yapıları soyutlayabilmeyi değil, fakat aynı zamanda bu soyut yapıları birbirine bağlayan bir nomolojik ağı zihninde canlandırabilmeyi gerektirir.

Önermeler ve Hipotezler

Figür 2.2 mesela zeka, çaba, akademik başarı ve kazanç poatansiyeli gibi teorik yapıların

birbirleriyle nasıl ilişkili olduklarını bir nomolojik ağda göstermektedir. Bu ilişkilerden her biri bir önerme olarak adlandırılır. Belirli bir fenomen veya davranışa ait açıklamalar bulmaya uğraşırken, yalnızca hedef fenomenin veya davranışın altında yatan kilit kavramları ve yapıları teşhis etmek kafi değildir. Aynı zamanda, bu yapılar arasındaki ilişkilerin biçimlerini teşhis etmek ve ifade etmek gereklidir. İlişkilerin böylesi biçimlerine önerme denir. Bir önerme yapılar arasında dekleratif bir formda ifade edilmiş tentatif ve konjektürel bir ilişkidir. Bir önerme misali şudur: “Öğrencinin zekasındaki bir artış onun akademik başarısında bir artışa neden olur.” Bu dekleratif ifade doğru olmak zorunda değil, fakat veri kullanılarak ampirik olarak test-edilebilir olmalıdır öyle ki doğru veya yanlış olup olmadığı yargısına varabilelim. Önergeler genellikle mantığa (tümdengelim) veya ampirik gözlemlere (tümevarım) dayanılarak türetilirler.

Önergeler doğrudan test edilemezler, çünkü soyut yapılar arasındaki ilişkilerdir. Onun yerine, yapılara tekabül eden ölçüler (değişkenler) arasındaki ilişkiyi incelemek suretiyle dolaylı olarak test edilirler. Bir önermenin ampirik formülasyonuna, değişkenler arasında bir ilişki olarak, **hipotez** denir (Figür 2.1). IQ skorları ve not ortalaması sırasıyla zekanın ve akademik başarının operasyonel ölçüleri oldukları için, yukarıdaki önerme hipotez formunda ifade edilebilir: “Öğrencinin IQ skorundaki bir artış onun not ortalamasında bir artışa neden olur.” Önergeler teorik düzlemde spesifikleştirilirken, hipotezler ampirik düzlemde spesifikleştirilirler. Böylece, hipotezler gözlem verisi kullanılarak ampirik olarak test-edilebilirler ve eğer ampirik gözlemlerce desteklenmezlerse reddedilebilirler. Elbette, hipotez testinin amacı mütakabil önermenin geçerli olup olmadığı çıkarımında bulunmaktır.

Hipotezler güçlü veya zayıf olabilirler. “Öğrencilerin IQ skorları akademik başarılarıyla ilişkilidir” hipotezi zayıf bir hipoteze misaldir, çünkü ne hipotezin yönselliğine (yani, ilişkinin pozitif veya negatif olup olmadığına) ve ne de nedenselliğine (yani, zekanın akademik başarıya veya akademik başarının zekaya neden olup olmadığına) işaret etmiyor. Daha güçlü bir hipotez “Öğrencilerin IQ skorları akademik başarılarıyla pozitif olarak ilişkilidir” hipotezidir ki yönselliğe işaret ediyor fakat nedenselliğe değil. Halihazırda, daha iyi bir hipotez ise “Öğrencilerin IQ skorlarının akademik başarılarına pozitif etkisi vardır” hipotezidir ki hem yönselliği ve hem de nedenselliği spesifikleştirir (yani, zeka akademik başarıya neden olur, tersi değil). Figür 2.2 deki işaretler ilgili hipotezlerin yönselliğini göstermektedir.

Keza ifade etmek gerekir ki bilimsel hipotezler açıklayıcı değişkenleri ve yanıt değişkenlerini açıkça spesifikleştirmelidirler. “Öğrencilerin IQ skorlarının akademik başarılarına pozitif etkisi vardır” hipotezinde zekanın açıklayıcı değişken (“neden”) ve akademik başarının yanıt değişkeni (“etki”) olduğu açıktır. Dahası, yine açıktır ki bu hipotez doğru (eğer daha yüksek zeka daha yüksek akademik başarıya yol açıyorsa) veya yanlış (eğer daha yüksek zeka daha yüksek akademik başarıya yol açmıyorsa ya da etki etmiyorsa) olarak değerlendirilebilir. Bu kitabın daha sonraki bölümlerinde, böylesi neden-ve-etki ilişkilerini ampirik olarak nasıl test edeceğimizi inceleyeceğiz. Mesela “öğrenciler genellikle zekidirler” veya “tüm öğrenciler akademik başarıya ulaşabilir” gibi ifadeler birer bilimsel hipotez değildir, çünkü açıklayıcı değişkenleri ve yanıt değişkenlerini spesifikleştirmiyorlar ve aynı zamanda doğru veya yanlış olarak değerlendirilebilecek bir yönsel ilişkiyi de spesifikleştirmiyorlar.

Teoriler ve Modeller

Bir **teori** ilgilenilen bir fenomeni veya bir davranışı belli sınırlayıcı koşullar ve varsayımlar altında açıklamaya ve tahmin etmeye matuf kendi aralarında sistematik olarak ilişkili yapılar ve önermelerin bir kümesidir. Esasen, bir teori birbiriyle bağlantılı teorik önermelerin sistematik bir topluluğudur. Önergeler genellikle iki veya üç yapı arasında bağlantı kurarken,

teoriler çoklu yapıların ve önermelerin bir *sistemini* temsil eder. Böylece, teoriler bu manada önermelere veya hipotezlere nazaran daha kompleks ve soyut ve daha geniş kapsamlıdır.

Şunu anmadan geçemeyeceğim ki bilimsel araştırmaya aşina olmayan insanlar bir teoriyi pek çok zaman bir *spekülasyon* olarak ya da *olgunun* (/fact) karşıtı olarak görürler. Mesela, şu sıkça dillendirilir ki öğretmenler ders verirken daha az teorik ve daha çok pratik veya olgusal (/factual) olmalılarıdır. Mamafih, pratik veya olgu teorisinin karşıtları değildir, fakat, bilimsel manada, bir teorisinin geçerliliğini test etmek için gereksinim duyulan esansiyel bileşenlerdir. İyi bir bilimsel teori, gözlenen olguları kullanmak suretiyle, iyice desteklenmiş olacaktır ve aynı zamanda pratik değeri olmalıdır; öte taraftan, kötü tanımlanmış bir teori bu boyutların eksikliğini yansıtacaktır. Meşhur organizasyonel psikoloji araştırmacısı Kurt Lewin (1890-1947), zamanında, şöyle demişti: “Pratik olmadan teori kısır; teori olmadan pratik kördür.” Bu yüzden, hem teori ve hem de olgular (veya pratik) bilimsel araştırma için esansiyeldir.

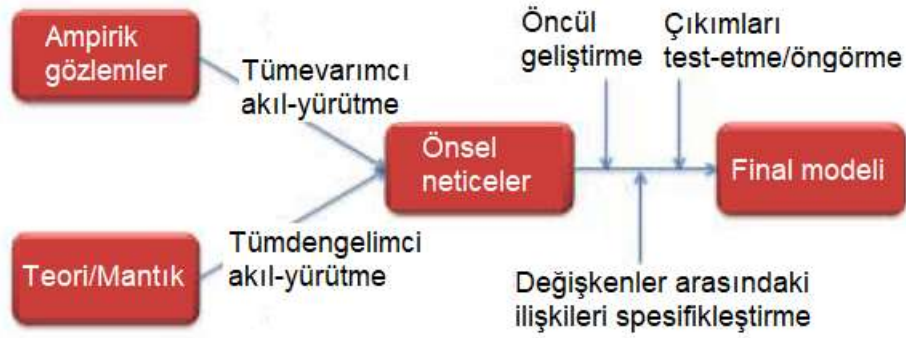
Teoriler sosyal veya doğal fenomenler için açıklamalar temin eder. Bölüm 1 de vurgulandığı üzere, bu açıklamalar iyi veya kötü olabilir. Böylece, iyi veya kötü teoriler olabilir. Bölüm 3 te bir teorisinin gerçekten nasıl iyi bir teori olduğunu değerlendirmekte kullanılabilecek bazı kriterleri betimleyeceğiz. Araştırmacıların şunu anlamaları önem arzeder ki teori “hakikat” değildir, herhangi bir teori hiç bir kutsallık taşımaz ve teoriler sırf birileri tarafından ileri sürüldüler diye kabul edilemezler. Bilimsel ilerleme boyunca, daha kötü teorilerin yerini, eninde sonunda, daha yüksek açıklayıcı gücü olan daha iyi teoriler alır. Araştırmacılar için asıl meydan okuma bir hedef fenomeni önceki teorilerden daha iyi açıklayabilecek daha iyi ve daha kavrayıcı teoriler inşa edebilmektir.

Teori ile beraber sıkça kullanılan bir terim model terimidir. Bir **model** bir sistemin tümsel veya kısmi bir temsildir ki o sistemi öğrenmeye-çalışmak için inşa edilir (mesela, sistem nasıl çalışır veya sistemi ne tetikler). Bir teori bir fenomeni açıklamaya çabalar, bir model bir fenomeni temsil etmeye çabalar. Modeller karar vericiler tarafından eldeki bir girdi kümesine dayanarak önemli kararlar vermek için sıkça kullanılır. Mesela, pazarlama müdürleri mesela önceki yılın reklam masrafları, satışlar, pazar büyümesi ve rakip ürünler gibi parametrelere dayanarak, farklı ürün hatlarının reklamına ne kadar para harcayacaklarına karar vermek için modelleri kullanabilirler. Benzerce, hava tahmincileri mesela rüzgar hızı, rüzgar yönü, sıcaklık ve nemlilik gibi parametrelere dayanarak, gelecekteki hava biçimini tahmin etmek için modelleri kullanabilirler. Bu modeller yararlı olmakla beraber, reklam giderini veya hava durumunu illaki açıklayamayabilirler. Modeller mesela matematiksel modeller, ağ modelleri ve iz modelleri gibi farklı tiplerde olabilirler. Modeller keza betimleyici, tahminci veya normatif olabilirler. Betimleyici modeller çoğu zaman kompleks sistemleri temsil etmek için kullanılırlar, bu sistemlerdeki değişkenleri ve ilişkileri zihinde canlandırmak gayesiyle. Bir reklam gideri modeli betimleyici bir model olabilir. Tahminci modeller (mesela, bir regresyon modeli) gelecekteki olayların öngörülmesine olanak sağlar. Hava tahmini modelleri tahminci modellerdir. Normatif modelleri müştereken kabul edilmiş normlar veya pratikler arasında kendi aktivitelerimize yön vermek için kullanırız. Modeller, aynı zamanda, eğer bir sistemin durumunu zamanın bir noktasında temsil ediyor ise, statik olabilirler; eğer bir sistemin zaman içerisinde evrimini temsil ediyor ise, dinamik olabilirler.

Teori veya model geliştirme prosesi tümevarımcı ve tümdengelimci akıl yürütmeyi içerebilir. Bölüm 1 den anımsanacağı üzere, **tümdengelim** bir fenomen veya davranış hakkında teorik veya mantıksal sebeplere ve bir ilk öncüller kümesine dayanarak neticelere varma prosesidir. Bir misal verirek, eğer belli bir banka çalışanlarını katı etik kurallara uymaya zorluyorsa (Öncül 1), ve Tankut bu bankanın bir çalışanı ise (Öncül 2), o zaman Tankut’un etik davrandığından emin olabiliriz (Netic). Tümdengelimde, eğer ilk öncüller ve sebepler yanlışsız ise neticelerin doğru olması gerekir.

Aksine, **tümevarım** olgulara veya gözlemlere dayanarak neticelere varma prosesidir. Mesela, eğer bir firma bir promosyon kampanyası için çok fazla para harcadıysa (Gözlem 1), ve buna rağmen satışlar artmadıysa (Gözlem 2), o zaman olanaklı ki promosyon kampanyası kötü yürütülmüştür (Netice). Mamafih, satışın kötülüğü için mesela ekonomik durgunluk veya rakip bir ürünün veya firmanın pazara girmesi veya belki tedarik zincirinde herhangi bir problem gibi alternatif açıklamalar olabilir. Tümevarımcı neticeler, bu yüzden, ancak bir hipotezdir ve aksi ispatlanabilir. Tümdengelimci neticeler tümevarımcı neticelere göre genellikle daha güçlü olma eğilimindedir, fakat yanlış bir öncüle dayanan tümdengelimci bir netice keza yanlıştır.

Figür 2.3 te gösterildiği üzere, tümevarımcı ve tümdengelimci akıl yürütme teoride ve model inşa etmede el ele gider. Tümevarım, bir olguyu gözlediğimiz ve “Bu niçin oluyor?” sorusunu sorduğumuz zaman vuku bulur. Bu soruyu cevaplamak suretiyle, bir veya daha çok sayıda tentatif açıklamayı ileri süreriz. Sonrasında, tentatif açıklamalar içerisinde mantığa ve makul öncüllere (yani, öğrenilmeye-çalışılan fenomene ait anlayışımıza) dayanarak en olası açıklamayı bulmak için tümdengelimci kullanırız. Araştırmacılar tümevarımcı ve tümdengelimci akıl yürütme arasında geri ve ileri doğru hareket edebilmelidir; eğer halihazırda mevcut bir modelde veya teoride genişletmeler veya modifikasyonlar yapacaklar ise, ya da onların daha iyilerini inşa edecekler ise, ki teoriler ve modeller bilimsel araştırmanın özüdür.



Figür 2.3 Model inşa etme prosesi

Bölüm 3

Araştırma Prosesi

Bölüm 1 de, gördük ki bilimsel araştırma bilimsel metodu kullanarak bilimsel bilgiye ulaşma prosesidir. Fakat böylesi bir araştırma nasıl yürütülür? Bu bölümde bilimsel araştırma prosesini ve araştırma prosesinin varsayımlarını (/assumptions) ve çıkımlarını (/outcomes) derinlemesine ele alacağız.

Sosyal Araştırmanın Paradigmaları

Araştırma dizaynımızı ve icramızı zihinsel referans modellerimiz veya çerçevelerimiz şekillendirir, ki onları akıl yürütmemizi ve gözlemlerimizi organize etmek için kullanırız. Bu zihinsel modellere veya çerçevelere (inanç sistemlerine) **paradigma** adı verilir. “Paradigma” sözcüğü Thomas Kuhn’un (1922-1996) *The Structure of Scientific Revolutions* (1962) kitabıyla ünlenmiştir ki o kitapta Kuhn bilimin ilerlemesini şekillendiren biçimleri ve aktiviteleri teşhis etmek için doğa bilimlerinin tarihini incelemiştir.¹ Benzer fikirler sosyal bilimlere de uygulanabilir, ki sosyal bilimlerde bir sosyal gerçekliğin farklı kişilerce farklı bakımlardan görülebilir olması, kişilerin gözlenen fenomen hakkında düşüncelerini ve akıl yürütmelerini sınırlandırabilir. Mesela, muhafazakarlar ve liberaller hükümetin (veya devletin) insanların yaşamındaki rolüne dair farklı algılara sahip olma ve bu yüzden sosyal problemlerin çözümüne dair farklı görüşlere sahip olma eğilimindedirler. Muhafazakarlar inanabilirler ki durgun ekonomiyi stimüle etmenin en iyi yolu vergileri düşürmektir, çünkü öyle yapmak insanların harcanabilir gelirini ve harcamayı artırır ve bu da dolayısıyla ticaret hacmini ve istihdamı artırır. Aksine, liberaller inanabilirler ki hükümetler daha doğrudan mesela kamu işleri ve altyapı projeleri gibi istihdam yaratan programlara yatırım yapmalıdır, çünkü öyle yapmak istihdamı ve insanların tüketim gücünü artırır ve ekonomiyi canlandırır. Benzerce, Batı toplumları mesela kişinin mahremiyet hakkı, özgür konuşma hakkı, silahlanma hakkı gibi bireysel haklara daha fazla önem atfederler. Aksine, Asya toplumları bireylerin haklarını ailelerin, organizasyonların ve hükümetlerin/devletlerin haklarına karşı dengeleme eğilimindedirler ve bu yüzden politikalarında daha çok müşterekçi ve daha az bireyseli olma eğilimindedirler. Böylesi perspektif farklılıkları çoğu zaman Batıların Asyalı hükümetleri/devletleri otokratik olmakla suçlamalarına ve öte taraftan Asyalıların da Batı toplumlarını açgözlülükle, yüksek suç oranlarına sahip olmakla ve bir “birey kültürü” yaratmakla eleştirmelerine yol açmaktadır. Bizim kişisel paradigmalarımız “renkli gözlüklere” benzer ki dünyayı nasıl gördüğümüze hükmeder ve dünyada gördüklerimiz hakkındaki düşüncelerimizi yapılandırır.

Paradigmaları tanımak çoğu zaman zordur, çünkü zımni, varsayılan ve farazilerdir. Mamafih, paradigmaları tanımak insanların aynı sosyal fenomeni algılamalarındaki farklılıkları anlamının ve bağdaştırmanın anahtarıdır. Mesela, liberaller orta öğretimi iyileştirmenin en iyi

¹ Çevirenin notu: Kuhn bilimsel bilginin ilerleyişine dair dikkat çekici iddiaların sahibidir. Şöyle ki bilim alanları doğrusal ve sürekli bir yolda ilerlemekten daha ziyade periyodik “paradigma değişikliklerine” uğrarlar ve bu paradigma değişiklikleri bilimciler için yeni yaklaşımların kapısını açar, öyleki daha önce asla geçerli olabileceğini mülhaza edemeyecekleri şeyleri anlayabilirler. Buna göre, bilimsel hakikat mefhumu zamanın herhangi bir belirli noktasında salt objektif kriterlerle tesis edilebilmez ancak bir bilimsel topluluk içerisinde bir konsensüsle tanımlanabilir. Rakip paradigmalar çoğunlukla kıyaslanabilir değildir: yani, gerçeğin birbiriyle rekabet eden ve telif-edilebilmez muhasebeleridir. Böylece, bizim bilimi şumullüce kavrayışımız asla bütünüyle tek başına “objektifliğe” dayanabilir. Bilim subjektif perspektifleri de hesaba katmalıdır. Çünkü tüm objektif neticeler nihayetinde araştırmacıların ve katılımcıların/deneklerin subjektif koşullanmaları/dünya-görüşleri üzerine bina edileceklerdir. (Wikipedia)

yolunun daha çok öğretmen işe almak olduğuna inanırken, niçin muhafazakarlar aynı amaca ulaşmak için öğretimi özelleştirmenin (mesela okul çekleri gibi vasıtalar kullanarak) daha efektif olduğuna inanır? Çünkü muhafazakarların rekabetçi piyasalara (yani, eğitime ayrılan para için okullar arasında oluşacak serbest rekabete) daha fazla inancı varken, liberaller daha fazla iş gücüne (yani, daha çok öğretmene ve okula sahip olmaya) inanır. Yine, sosyal bilim araştırmasında, belli bir teknolojiyi bir organizasyon başarıyla kullanırken, niçin bir başka organizasyonun felaket başarısız olduğunu anlamak için, dünyaya “rasyonel mercekten” bakan bir araştırmacı, probleme yönelik mesela kifayetsiz teknoloji veya teknoloji ve o teknolojinin kullanıldığı iş bağlamı arasındaki uyum eksikliği gibi rasyonel açıklamalar bulmaya çalışacaktır. Aynı probleme “sosyal mercekten” bakan diğer bir araştırma, kifayetsiz kullanıcı eğitimi veya yönetim desteği eksikliği gibi sosyal kusurlar ararken, “politik mercekten” bakanlar, teknoloji tatbikatı prosesini akamete uğratabilecek organizasyonel politikaların mevcudiyeti üzerinde durur. Bu sebeple, bilinçaltı paradigmlar, araştırmacıların ölçmek istedikleri kavramları, gözlemlerini ve bir fenomene dair müteakip interpretasyonlarını kısıtlatırlar. Mamafih, sosyal fenomenin karmaşık doğası göz önüne alındığında, yukarıdaki paradigmların hepsinin kısmen doğru olması mümkündür ve problemin tam anlaşılması çoklu paradigmların anlaşılmasını ve uygulanmasını gerektirebilir.

Bugün, sosyal bilim araştırmacıları arasında popüler olan iki paradigma, pozitivism ve post-pozitivismdir. Fransız filozofu Auguste Comte'un (1798-1857) çalışmalarına dayanan **pozitivism**, 20. yüzyılın ortalarına kadar baskın bir bilimsel paradigmaydı. Bilim veya bilgi yaratımının gözlenebilecek ve ölçülebilecek olanla sınırlandırılması gerektiğini savunur. Pozitivism, yalnızca doğrudan test edilebilecek teorilere dayanma eğilimindedir. Pozitivism başlangıçta bilimsel tahkikatı dinden (ki orada akaid nesnel olarak gözlenemez) ayırmak için bir çaba olsa dahi, pozitivism ampirizme veya gözlenen veriye körü körüne inanmaya ve gözlenebilir olguların ötesine uzanma veya akıl yürütme çabasının reddedilmesine yol açmıştır. İnsan düşünceleri ve duyguları doğrudan ölçülebilmediğinden, bilimsel araştırma için meşru mevzular olarak telakki edilmemişleridir. Pozitivist felsefenin katı ampirik doğasına dair uğranılan hüsranslar, 20. yüzyılın ortalarında **post-pozitivizmin** (ya da postmodernizmin) gelişimine yol açtı. Post-pozitivism, ampirik gözlemleri mantıksal akıl yürütmeyeyle birleştirerek, bir fenomen hakkında makul çıkarımlarda bulunulabileceğini savunur. Post-pozitivistler, bilimi kesin değil ama olasılıksal (yani, pek çok olumsuzluğa (*/contingency*) dayanan) olarak görürler ve genellikle sosyal gerçekliği daha iyi anlamak için bu olumsuzlukları keşfetmeye çalışırlar. Post pozitivistler kampı, kendi aralarında da ikiye ayrılmışlardır: subjektivistler ve eleştirel realistler. *Subjektivistler*, dünyayı nesnel bir gerçeklikten ziyade subjektif aklımızın subjektif bir inşası olarak görürken, *eleştirel realistler*, kişinin düşüncesinden bağımsız bir dışsal gerçek olduğuna ve böyle bir gerçeğin herhangi bir derecede kesinlikle asla bilinebilmeyeceğine inanırlar.

Burrell ve Morgan (1979), *Sociological Paradigms and Organizational Analysis* adlı seminal kitaplarında ileri sürdüler ki sosyal bilim araştırmacılarının sosyal fenomenleri görme ve öğrenmeye-çalışma yolunu felsefi varsayımların iki temel kümesi şekillendirir: ontoloji ve epistemoloji. **Ontoloji**, dünyayı nasıl gördüğümüz hakkındaki varsayımlarımıza dairdir; mesela, dünyanın çoğunlukla sosyal düzenden mi yoksa sürekli değişimden mi ibaret olduğu. **Epistemoloji**, dünyayı öğrenmeye-çalışmanın en iyi yolu hakkındaki varsayımlarımıza dairdir; mesela, sosyal gerçeği incelemek için objektif mi yoksa subjektif bir yaklaşım mı kullanmalıyız. Bu iki varsayımlar kümesini kullanarak, sosyal bilim araştırmasını dört kategoriden birine ait olmak üzere tasnif edebiliriz (bkz. Figür 3.1).

Eğer araştırmacılar, dünyayı çoğunlukla sosyal düzenden (ontoloji) ibaret olarak görüyor ve böylece, düzenli olayların veya davranışların kalıplarını öğrenmeye-çalışıyor ve böylesi bir dünyayı öğrenmeye-çalışmanın en iyi yolunun objektif, yani, yoklama benzeri

standartlaştırılmış veri toplama araçlarını kullanamamak gibi, gözlemi veya interpretasyonu, yapan kişiden bağımsız kılan bir yaklaşım kullanmak (epistemoloji) olduğuna inanıyorlar ise, o zaman bir **fonksiyonalizm** paradigması benimsemiş oluyorlar. Mamafih, eğer sosyal düzeni öğrenmeye-çalışmanın en iyi yolunun, aksine, mesela farklı katılımcılarla görüşmek ve kendi subjektif perspektiflerini kullanarak verilen yanıtlar arasındaki farklılıkları uyuşturmakta (*reconciling*) olduğu gibi, içerilen katılımcıların subjektif interpretasyonu olduğuna inanıyorlar ise, o zaman bir **interpretivizm** paradigması kullanmaktadırlar. Eğer araştırmacılar dünyanın radikal bir değişimden ibaret olduğuna inanıyor ve objektivist bir yaklaşım kullanarak değişimi anlamaya çalışıyorlar ise, o zaman bir **radikal strüktüralizm** paradigması kullanıyorlar demektir. Eğer sosyal değişimi, içerilen katılımcıların subjektif perspektiflerini kullanarak anlamak isterlerse, o zaman, araştırmacılar, bir **radikal hümanizm** paradigması takip ediyor olurlar.



Figür 3.1 Sosyal bilim araştırmasının dört paradigması (Burrell ve Morgan, 1979)

Bugüne kadar, sosyal bilim araştırmalarının çoğu doğal bilimleri taklit etmiş ve fonksiyonalist paradigmayı takip etmiştir. Fonksiyonalistler, sosyal düzen ya da kalıpların fonksiyonel bileşenleri cinsinden anlaşılabilirliğine inanırlar ve bu yüzden bir problemi küçük bileşenlere ayırmaya ve yoklamalar ve deneysel araştırma gibi objektivist teknikleri kullanarak bir ya da daha fazla bileşeni ayrıntılı olarak incelemeye çalışırlar. Ancak, post-positivist düşünmenin ortaya çıkmasıyla birlikte, şimdilik az ama giderek artan sayıda sosyal bilim araştırmacıları, sosyal düzeni, röportajlar ve etnografik öğrenme-çalışmaları gibi subjektivist teknikleri kullanarak anlamaya çaba göstermektedir. Radikal hümanizm ve radikal strüktüralizm, sosyal bilim araştırmalarının göz ardı edilebilir bir orantısını temsil etmeye devam etmektedir, çünkü bilim adamları, nev-i şahsına münhasır ve değişen olaylar yerine, davranışın, olayların veya fenomenlerin genelleştirilebilir kalıplarını anlamakla başlıca ilgilenmektedirler. Yine de, sosyal değişimi incelemek istiyorsan, mesela Ortadoğu ülkelerinde demokratik hareketlerin niçin giderek artmakta olduğu, ya da bu hareketin niçin Tunus'ta başarılı olduğu, Libya'da başarıya ulaşmasının daha uzun bir zaman aldığı ve Suriye'de hala başarıya ulaşamadığı gibi, o zaman, belki de radikal hümanizm böylesi bir öğrenme-çalışması için doğru yaklaşımdır. Sosyal ve organizasyonel fenomenler genellikle hem düzenin hem de değişimin unsurlarını içerir. Mesela, organizasyonel başarı, formalize edilmiş işletme proseslerine, çalışma prosedürlerine ve vazife sorumluluklarına bağlı olmakla beraber; işletme ortamında yer alan rakipler, rakip ürünler, tedarikçiler ve müşteri tabanının sürekli olarak değişen bir karışımı tarafından da eş zamanlı olarak kısıtlanır. Bu yüzden, mesela bazı organizasyonların neden diğerlerinden daha başarılı

olduğu gibi sosyal fenomenleri bütünsel ve daha eksiksiz olarak anlamak, sosyal araştırmaya çoklu-paradigmatik bir yaklaşımın takdirini ve uygulanmasını gerektirir.

Araştırım Prosesini Gözden-geçirme

Peki, zihinsel paradigmalarımız, sosyal bilim araştırmasını nasıl şekillendiriyor? Özünde, tüm bilimsel araştırmalar gözlemi, rasyonelleştirmeyi ve geçerlemeyi (/validation) içeren bir iteratif prosestir. **Gözlem** fazında, bizi ilgilendiren doğal veya sosyal bir fenomeni, olayı veya davranışı gözleriz. **Rasyonelleştirme** fazında, gözlediğimiz fenomeni, olayı veya davranışı, elde edilen gözlemler bulmacasının farklı parçalarını mantıksal olarak birbirine bağlamak suretiyle manalandırmaya çabalarız ki bu bazı durumlarda bir teorinin inşasına yol açabilir. Son olarak, **geçerleme** fazında, teorilerimizi, bir veri toplama ve analizi prosesi boyunca ilerleyen bir bilimsel metot kullanmak suretiyle test ederiz; ve bunu yaparken, başlangıç teorimizi modifiye edebilir ya da genişletebiliriz. Mamafih, araştırım dizaynları, araştırmacının gözlem yapmakla başlayıp, ondan sonra gözlemlerini rasyonelleştirmeye mi çabaladığına (tümevarımcı araştırım) yoksa araştırmacının ex ante bir realizasyon veya bir teoriyle başlayıp ondan sonra o teoriyi geçerlemeye mi çabaladığına (tumdengelimci araştırım) bağlı olarak değişiklik gösterir. Bu sebeple, gözlem-rasyonelleştirme-geçerleme döngüsü, Bölüm 1 de tartışılan araştırmanın tümevarım-tumdengelim döngüsüne çok benzer.

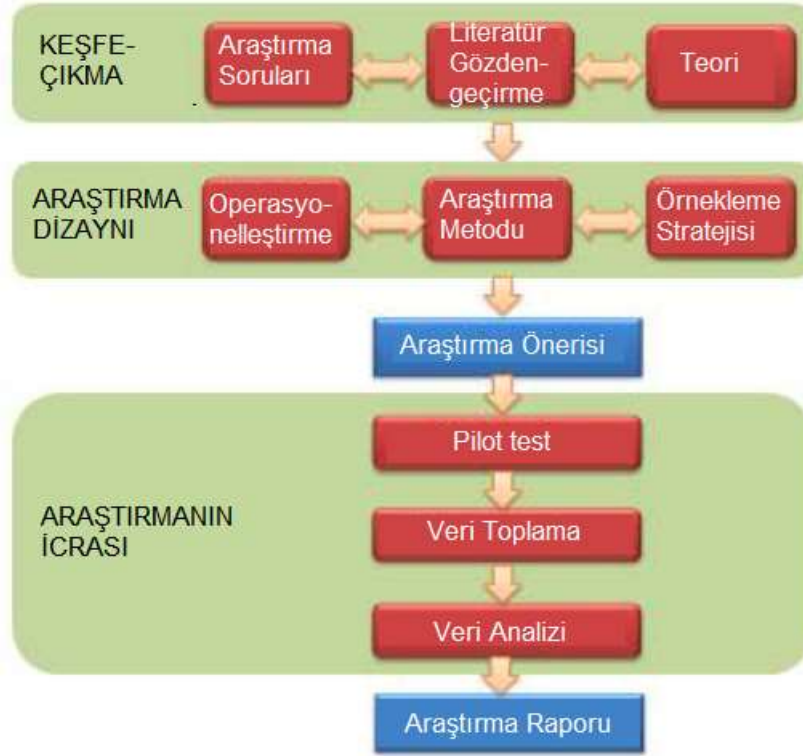
Çoğu geleneksel araştırım, doğası itibarıyla tumdengelimci ve fonksiyonalistik olma eğilimindedir. Şekil 3.2 böyle bir araştırım projesinin şematik bir görünümünü sunmaktadır. Bu figür, keşfe çıkma (/exploration), araştırım dizaynı ve araştırmanın icrası olmak üzere üç fazda kategorize edilen fonksiyonalist araştırmalarda gerçekleştirilecek bir dizi aktiviteyi resmediyor. Bu genelleştirilmiş dizaynın tüm araştırmalar için bir yol haritası veya akış şeması olmadığını unutmamak gerek. Bu, sadece fonksiyonalistik araştırmalara uygulanır ve spesifik bir projenin gereksinimlerini karşılamak üzere modifiye edilebilir ve edilmelidir.

Araştırmanın ilk fazı, **keşfe çıkmaktır**. Bu faz, soruşturmayı derinleştirecek araştırım sorularını keşfetmeye çalışmayı ve seçmeyi, tahkikat alanında halihazırda mevcut olan bilgiyi anlamak için, o alanda yayımlanmış literatürü incelemeyi ve ilgilenilen araştırım sorularını cevaplamaya yardımcı olabilecek teorilerin ifade edilmesini ihtiva eder.

Keşfe çıkma fazında ilk adım, spesifik bir davranış, olay veya ilgilenilen fenomene dair bir veya daha çok **araştırım sorusunu** ifade etmektir. Araştırım soruları, bir davranış, olay veya ilgilendiğin fenomen hakkında spesifik sorulardır ki araştırmada cevaplarını bulmak istersin. Misal olarak, satıcılarını tanımaksızın, tüketicileri internet üzerinden mal veya hizmet satın almaya motive eden faktörlerin hangileri olduğu, lise öğrencilerini nasıl daha yaratıcı hale getirebileceğimiz ve bazı kişilerin niçin terörist eylemler gerçekleştirdiği verilebilir. Araştırım soruları ne, neden, nasıl, ne zaman ve benzeri soru kelimeleri ile kurularak, meselenin derinine inilmeye çabalanır. Daha ilginç araştırım soruları şunlardır ki daha geniş bir popülasyonu hedef alırlar (mesela, “firmalar nasıl inovatif olabilir” sorusu, “Çinli firmalar hizmet sektöründe nasıl inovatif olabilir” sorusuna göre daha ilginç bir araştırım sorusudur), gerçek ve karmaşık problemleri ele alırlar (hipotetik veya basit problemlerin haricinde) ve cevapları aşikar değildir. Dar odaklı araştırım soruları (genellikle ikili bir evet/hayır cevabı olan), daha az yararlı ve daha az ilginç olmak durumunda kalırlar ve sosyal fenomenlerin ince nüanslarını yakalamak için daha az uygunlardır. İlginç olmayan araştırım soruları genellikle ilginç olmayan ve yayımlanabilir araştırım bulgularına yol açarlar.

Bir sonraki adım, ilgilenilen alanda bir **literatür gözden-geçirmesi** yapmaktır. Literatür gözden-geçirmesi üç maksada ulaşmak üzere yapılır: (1) tahkikat alanındaki mevcut bilgiyi yoklamak, (2) o alandaki önemli yazarları, makaleleri, teorileri ve bulguları teşhis etmek ve (3)

araştırma alanındaki bilgi boşluklarını teşhis etmek. Literatür gözden-geçirmesi, günümüzde, internet veritabanlarında, arama motorlarına anahtar kelimeleri girmek suretiyle kolayca yapılabilmektedir. Anahtar kelimeler, arama sonuçlarını daraltmak veya genişletmek için “ve” ve “veya” operatörleri kullanılarak birleştirilebilir. Anahtar kelimeler ile yapılan aramada, alakalı makalelerin kısa bir listesi oluşturulduktan sonra, araştırmacı her makaleyi veya en azından özet bölümünü münhasıran tarayarak, o makalenin ayrıntılı bir gözden geçirme için uygunluğunu belirlemelidir. Literatür gözden-geçirmeleri adanmaklılı tamamlanmalı ve birkaç dergi, birkaç yıl veya spesifik bir metodoloji ile sınırlı olmamalıdır. Gözden geçirilen makaleler, tablolar şeklinde özetlenebilir ve kavram matrisi gibi organizasyon çerçeveleri kullanılarak daha ileri düzeyde yapılandırılabilirler. İyi yapılmış bir literatür gözden-geçirmesi, ilk araştırma sorularının literatürde daha önce ele alınıp yazılmadığını (ki bunları tekraren öğrenmeye-çalışma ihtiyacını ortadan kaldırır), yeni veya daha ilginç araştırma soruları olup olmadığını ve orijinal araştırma sorularının literatür gözden-geçirmesi bulguları ışığında modifiye edilmesinin veya değiştirilmesinin gerekip gerekmediğini göstermelidir. Gözden geçirme, ayrıca, ilgilenilen sorulara dair bazı sezgiler veya potansiyel cevaplar sağlayabilir ve/veya daha önce benzer soruları cevaplamak için kullanılmış olan teorileri teşhis etmeye yardımcı olabilir.



Figür 3.2 Fonksiyonalistik araştırma prosesi

Fonksiyonalist (tumdengelimci) araştırma, teorilerin test edilmesini ihtiva ettiği için, üçüncü adım, arzulanan araştırma sorularına cevap vermeye yardımcı olabilecek bir veya birden çok **teoriyi** ifade etmektedir. Literatür gözden-geçirmesi, ilgilenilen fenomene ilişkin geniş bir yelpazedeki potansiyel kavramları veya yapıları ortaya çıkarabilirken, bir teori, bu yapılardan hangilerinin hedef fenomenle mantıklıca alakalı olduğunu ve nasıl alakalı olduğunu belirlemeye yardımcı olacaktır. Desteklenmeyen teoriler, geniş bir yelpazede daha az alakalı, ucu ucuna alakalı veya alakasız yapıları ölçmekle beraber, sadece şans eseri olmayan ve manalı sonuçların elde edilme şansını da minimize etmekle de sonuçlanabilir. Fonksiyonalist araştırmada, teoriler ampirik testleri yapmak için hipotezlerin postülasında mantıksal temel

olarak kullanılabilir. Aşıkâr ki tüm teoriler tüm sosyal fenomenleri öğrenmeye-çalışmak için uygun değildir. Teoriler, hedef problemle uyuşmalarına ve varsayımlarının hedef problem ile ne kadar tutarlı olduğuna bağli olarak dikkatli bir şekilde seçilmelidir. Bir sonraki bölümde teorileri ve teorileştirme prosesini ayrıntılı olarak inceleyeceğiz.

Araştırım prosesinin bir sonraki fazı, **araştırım dizaynıdır**. Bu proses, keşfe çıkma fazında ifade edilen araştırım sorularına tatmin edici cevaplar vermek için yapılacak aktivitelerin detaylı bir planını çıkarmayı ihtiva eder. Bir araştırım metodunun seçilmesi, ilgilenilen yapıların operasyonelleştirilmesi ve mütenasip bir örnekleme stratejisinin belirlenmesi bu aktiviteler arasında yer alır.

Operasyonelleştirme, soyut teorik yapılar için kesin ölçülerin dizayn edilmesi prosesidir. Bu, sosyal bilim araştırımında büyük bir sorundur, şöyle ki mesela önyargı, yabancılaştırma ve liberalizm gibi pek çok yapının tanımlanması zaten zor iken, bir de bunları tek başına isabetlice ölçmek çok daha zordur. Operasyonelleştirme, ilgilenilen yapıların bir “operasyonel tanımının” (veya “kavramsallaştırmasının”) spesifikleştirilmesi ile başlar. Ardından, araştırmacı, kendi ilgilendiğı yapıları ölçmek için doğrudan veya modifiye ederek kullanabileceğı ve operasyonel tanımlarıyla eşleşen önceden geçerlenmiş ölçülerin mevcut olup olmadığını görmek üzere literatürde arama yapabilir. Eğer böyle ölçüler mevcut değilse veya mevcut ölçüler yetersizse veya araştırmacının düşündüğünden farklı bir kavramsallaştırmayı yansıtıyorlarsa, ilgilenilen yapıların ölçülmesi için yeni enstrümanların dizayn edilmesi gerekebilir. Bu, arzu edilen yapının tam olarak nasıl ölçüleceğini spesifikleştirmek manasına gelir (mesela, kaç madde/soru, hangi maddeler/sorular ve benzeri). Çoğu zaman, bu proses uzun ve zahmetli olabilir, çünkü yeni dizayn edilen enstrümanın “bilimsel olarak geçerli” kabul edilebilmesinden önce, çok sayıda öntest ve modifikasyon aşamalarından geçmesi gerekir. Ölçüm hakkında sonraki bir bölümde, yapıların operasyonelleştirilmesini tartışacağız.

Operasyonelleştirmeyle eşzamanlı olarak, araştırmacı ayrıca, araştırım sorularına cevap vermek üzere veri toplamak için hangi **araştırım metodunu** kullanmak istediğine de karar vermelidir. Bu metotlar, mesela deneyler veya yoklama araştırması gibi nicel metotlar ya da mesela vaka araştırması veya eylem araştırması gibi nitel metotlar ya da mümkün ki her ikisinin bir kombinasyonu olabilir. Eğer bir deney arzu ediliyor ise, o zaman deney dizaynı nedir? Eğer bir yoklama arzu ediliyor ise, bir posta yoklaması mı, telefon yoklaması mı, web yoklaması mı yoksa bunların bir kombinasyonunu mu planlıyorsun? Kompleks, bellisiz ve çok yönlü sosyal fenomenler için, çok metotlu yaklaşımlar belki daha uygun olabilir; böylece, hem her bir araştırım metodunun kendine özgü güçlü yanlarından faydalanılabilir ve hem de tek bir metot kullanılarak elde edilemeyecek kavrayışlara sahip olunabilir.

Araştırmacılar ayrıca, veri toplamak istedikleri hedef popülasyonu ve o popülasyondan bir örnek seçmek için uygulayacakları **örnekleme** stratejisini dikkatle seçmelidir. Mesela, bireyleri mi yoksa firmaları mı yoksa firmalardaki çalışma gruplarını mı yoklayacaksın? Hangi tip bireyleri ya da firmaları hedeflemek istiyorsun? Örnekleme stratejisi, bir araştırım probleminin analiz ünitesi/birimi ile yakından ilişkilidir. Bir örnek seçerken, yanlış gözlemler üretebilecek yanlış bir örnekten (mesela, elverişliliğe dayanan örnek) kaçınmak için, gereken özen gösterilmelidir. Daha sonraki bir bölümde örnekleme derinlemesine ele alınmaktadır.

Bu safhada, araştırım prosesinin önceki safhalarında verilen tüm kararları ve her kararın ardındaki mantığı detaylandıran bir **araştırım önerisi** yazmak genellikle iyi bir fikirdir. Bu çok bölümlü öneri, araştırmacının hangi araştırım sorularını ve niçin öğrenmeye-çalışmayı arzu ettiğini, bu alanda önceki bilgi durumunu, test edilecek hipotezlerle beraber istihdam etmek istediğı teorileri, yapıları nasıl ölçeceğini, hangi araştırım metodunu ve niçin kullanacağını ve arzulanan örnekleme stratejisini dillendirmelidir. Araştırmalara fon temin eden kurumlar, genellikle, aralarından en iyisini seçmek üzere böyle bir öneri isterler. Bir araştırım projesi için

fon bulma kaygısı yoksa bile, bir öneri, veri toplamaya başlamadan önce, başka araştırmacıların görüşlerini almak ve araştırma projesinde potansiyel problemleri (mesela, bazı önemli yapıların öğrenme-çalışmasında eksik olup olmadığı gibi) teşhis etmek için yararlı bir araç olarak kullanılabilir. Bu başlangıç görüşleri paha biçilemezdir, çünkü bir araştırma öğrenme-çalışmasında veri toplandıktan sonra farkedilen kritik problemleri ortadan kaldırmak için çoğu zaman çok geç kalınmış olur.

Kimi öğrenmeye-çalışacağına (öznel/bireyler), neyi ölçeceğine (kavramlar) ve veriyi nasıl toplayacağına (araştırma metodu) karar verdikten sonra, araştırmacı artık **araştırmanın icrası** fazına geçmeye hazırdır.

Pilot testi genellikle araştırma prosesinin gözden kaçan fakat aşırı derece önemli bir parçasıdır. Bir pilot testi yapmak, araştırma dizaynı ve/veya enstrümantasyondaki potansiyel problemleri (mesela, sorulan soruların hedeflenen örnek için anlaşılabilir olup olmadığı gibi) tespit etmemize yardımcı olur ve öğrenme-çalışmasında kullanılan ölçüm enstrümanlarının, ilgilenilen yapıların güvenilir ve geçerli ölçüleri olduğundan emin olmamızı sağlar. Pilot örnek genellikle hedef popülasyonun küçük bir alt kümesidir. Başarılı bir pilot testinden sonra, araştırmacı örneklediği popülasyonun üzerine **veri toplamaya** devam edebilir. Toplanan veri, istihdam edilen araştırma metoduna bağlı olarak nicel veya nitel olabilir.

Veri toplamayı müteakip, veri analiz edilir ve ilgilenilen araştırma sorularına ilişkin neticeler çıkarmak maksadıyla interpretasyonu yapılır. Toplanan verinin tipine (nicel veya nitel) bağlı olarak, **veri analizi** nicel (mesela, regresyon veya strüktürel eşitlik modellemesi gibi) veya nitel (mesela, kodlama veya içerik analizi gibi) olabilir.

Araştırmanın son fazı, nihai **araştırma raporunun** hazırlanmasını da ihtiva eder ki bu rapor, baştan sona araştırma prosesini ve tüm bulgularını bir araştırma makalesi, tez veya monograf biçiminde belgeler. Bu raporda, araştırma prosesi dahilinde yapılan tüm seçimler (mesela, kullanılan teori, seçilen yapılar, kullanılan ölçüler, araştırma metotları, örnekleme ve benzeri) ve niçin yapıldıklarının yanısıra araştırma prosesinin her bir fazının çıkımları belli bir çerçevede ayrıntılı olarak yazılmalıdır. Araştırma prosesi öyle ayrıntılı betimlenmelidir ki başka araştırmacıların bizim öğrenme-çalışmamızı yinelemesine, bulgularımızı test etmesine ya da yaptığımız çıkarımların bilimsel olarak kabul edilebilir olup olmadığına karar vermesine olanak sağlasın. Elbette, halihazırda yazılmış bir araştırma önerisine sahip olmak, nihai raporun yazılma prosesini hayli basitleştirecek ve hızlandıracaktır. Aşık ki araştırmanın bir değeri olmayacaktır, eğer araştırma prosesi ve çıkımları meslektaşların eleştirisi ve gelecek nesiller için belgelenmez ise. Belgeleme, bilimin ilerlemesi için gereklidir.

Araştırmalarda Yapılan Müşterek Yanlışlıklar

Araştırma prosesi, problemler ve tuzaklar ile doludur ve acemi araştırmacılar, bir araştırma projesine hatırı sayılır miktarda zaman ve çaba harcadıktan sonra, çoğu zaman, araştırma sorularının yeterince cevaplanmadığını veya bulguların yeterince ilginç olmadığını veya araştırmanın “kabul edilebilir” bilimsel nitelikte olmadığını görebilirler. Bu tür problemler genellikle araştırma makalelerinin dergiler tarafından reddedilmesine neden olur. Sıkça yapılan bazı yanlışlıklar aşağıda betimlenmektedir.

Motivasyonu yetersiz araştırma soruları. Çoğu zaman, “elimizin altındaki” problemleri seçeriz ki bize ilginç geliyor olabilirler fakat bilimsel topluluğun geneline değil; yani, bunlar soruşturulan fenomen hakkında yeni bir bilgi ya da içgörü üretmez. Araştırma prosesi, araştırmacı tarafında önemli bir zaman ve çaba yatırımı demek olduğuna göre, araştırmacı, cevaplamaya çalıştığı araştırma sorularının sahiden bir popülasyonun önemli bir bölümünü

etkileyen ve önceki arařtırmalarda kafi derecede ele alınmamıř gerek problemlere (ve hipotetik problemlere deęil) deęindięinden emin olmalıdır (ve bařkalarını ikna edebilmelidir).

Arařtırma modasına uymak. Dięer bir müřterek yanlıřlık, sınırlı raf ömrüne sahip “popüler” mevzuları takip etmektir. Tipik bir misal, günümüzde popüler olan teknolojileri veya uygulamaları öğrenmeye-alıřmaktır. Arařtırmanın tamamlanması ve yayımlanması birkaç yıl süreceęi için, moda olan mevzuya yönelik popüler ilgi, arařtırmanın tamamlandıęı ve yayımlanmak üzere sunulduęu sırada azalabilecek veya yok olabilecektir. Daha iyi bir strateji, önemini yıllar boyunca sürdüren “zamansız” mevzuları öğrenmeye-alıřmak olabilir.

Arařtırılabilmeyen problemler. Bazı arařtırma problemlerine, yalnızca gözlenen kanıta dayanarak veya halihazırda kabul gören metot ve prosedürleri kullanarak kafi derecede özüm bulunamayabilir. Bu tür problemlerden kaçınmak en iyisidir. Mamafih, bazı arařtırılabilmeyen, muęlak tanımlanmıř problemler modifiye edilmek veya uyarlanmak suretiyle, iyi tanımlanmıř ve yararlı problemlere dönüřtürülebilirler.

Favori arařtırma metotları. Pek ok arařtırmacı, belli bir arařtırma problemini yeniden ele alma eęilimindedir, ünkü ařına oldukları arařtırma metotlarını (mesela, yoklama arařtırması gibi) kolayca uygulamaya elverişlidir. Bu talihsiz bir eęilim elbette. Arařtırma metotları, bir arařtırma problemine en iyi uyumu göstermesi bakımından seilmelidir, ve bařka bir řekilde deęil.

Kör veri madencilięi. Bazı arařtırmacılar önce veri toplama eęilimindedirler (halihazırda mevcut olan enstrümanları kullanarak) ve daha sonra topladıkları veri ile ne yapacaklarını anlamaya alıřırlar. řunu unutma ki veri toplama, arařtırmanın planlamasını, dizayn edilmesini ve icra edilmesini kapsayan uzun ve tafsilatlı bir prosesin yalnızca bir adımıdır. Olguda, veri toplamının öncesinde, bir arařtırma prosesinde, bir dizi bařka faaliyete ihtiya vardır. Arařtırmacılar, ayrıntılı bir planlama yapmadan, veri toplamaya kalkarlarsa, toplanan veri büyük olasılıkla alakasız, kusurlu veya yararsız olacak ve veri toplama abaları tamamen bořa ıkabilecektir. Verinin bolluęu, arařtırmanın planlamasındaki ve dizayn edilmesindeki kusurları asla gideremez ve bilhassa ilgi uyandıracak bir arařtırma sorusunun eksiklięini telafi edemez.

Bölüm 4

Bilimsel Araştırmada Teoriler

Daha önceki bölümlerden öğrendiğimiz üzere, bilim, bilimsel metot kullanılarak türetilen bir “teoriler” koleksiyonu olarak sunulan bilgidir. Bu bölümde, bir teorinin ne olduğunu, niçin araştırmada teorilere ihtiyaç duyduğumuzu, bir teorinin unsurlarının neler olduğunu, teorileri nasıl değerlendireceğimizi, araştırmada teorileri nasıl uygulayabileceğimizi inceleyeceğiz ve beraberinde sosyal bilim araştırmasında sıklıkla kullanılan beş teoriye değinerek misallendireceğiz.

Teoriler

Teoriler, doğal veya sosyal bir davranış, olay veya fenomenin açıklamalarıdır. Daha formal olarak söylemek gerekirse, bir bilimsel teori, bir yapılar (kavramlar) ve önermeler (o yapılar arasındaki ilişkiler) sistemidir ki bu sistem ilgilenilen fenomenin, bazı varsayımlar (*/assumptions*) ve sınırlandırıcı koşullar altında mantıksal, sistematik ve bağdaşık bir açıklamasını kollektif olarak ortaya koyar (Bacharach 1989).²

Teoriler, bir şeyi yalnızca betimlemek veya tahmin etmekten ziyade, niçin olduğunu açıklamalıdır. Dikkatini çekmek isterim ki olayları veya davranışları, niçin meydana geldiklerini açıklamak zorunluluğu olmaksızın, bir tahminciler kümesi kullanarak tahmin etmek olanaklıdır. Mesela, piyasa analistleri, borsadaki dalgalanmaları tahmin etmek için, daha önce gözlenen korelasyonlar ışığında, piyasa duyuruları, büyük şirketlerin kazanç raporları ve ilgili ajansların temin ettiği taze veriyi baz alır. Tahmin için gereken, sadece korelasyonlardır. Aksine, açıklamalar *nedensellikleri* ya da neden-ve-etki ilişkilerini anlamayı gerektirir. Nedenselliğin tesisi için üç koşul zaruridir: (1) iki yapı arasındaki korelasyonlar, (2) zamansal öncelik (neden, zaman içinde etkiden önce gelmelidir) ve (3) alternatif hipotezlerin reddedilmesi (test yoluyla). Bilimsel teoriler, teolojik, felsefi veya başka açıklamalardan farklıdır, şöyle ki bilimsel teoriler bilimsel metotlar kullanılarak ampirik olarak test edilebilir.

Açıklamalar, idiografik veya nomotetik olabilir. **İdiografik açıklama**, tek bir durumu veya olayı kendine özgü (*/idiosyncratic*) detayıyla açıklar. Mesela, bir sınavda başarısız oldun çünkü: (1) o gün bir sınavın olduğunu unutmuştun, (2) trafik sıkışıklığından dolayı sınava geç kaldın, (3) sınav sırasında panik yaptın, (4) bir önceki akşam geç saatlere kadar işyerinde çalışmak zorunda kaldığın için, sınava ders-çalışmadın, ya da hatta (5) köpeğin ders kitabını yedi. Bu açıklamalar detaylı, isabetli (*/accurate*) ve geçerli (*/valid*) olabilirler, ancak başka benzer durumlar için, hatta aynı kişi için bile, uygulanamayabilirler ve bu yüzden de genellenebilir değillerdir. Aksine, **nomotetik açıklamalar**, spesifik bir durum veya olaydan ziyade, bir durumlar veya olaylar sınıfını açıklamaya çalışır. Mesela, öğrenciler sınavlardan zayıf alır, çünkü sınavlara hazırlanmak için harcadıkları zaman kafi gelmez veya gerginlik, dikkat eksikliği veya başka bir tıbbi rahatsızlığa maruz kalırlar. Nomotetik açıklamalar durumlar, olaylar veya insanlar arasında genellenebilir olmak üzere dizayn edildikleri için, daha az kesin (*/precise*), daha az tam ve daha az detaylı olma eğilimindedirler. Mamafih, yalnızca birkaç açıklayıcı değişken kullandıkları için, açıklamaları daha ekonomiktir. Maksat, teorileri, aynı zamanda, davranışların veya fenomenlerin biçimleri için genelleştirilmiş açıklamalar olarak

² Bacharach, S. B. (1989). “Organizational Theories: Some Criteria for Evaluation”, *Academy of Management Review* (14;4), 496-515.

ortaya koymak olduğuna göre, teorik açıklamalar doğası gereği genellikle nomotetiktir.

Teorinin ne olduğunu öğrenirken, aynı zamanda teorinin ne olmadığını öğrenmek de önemlidir. Teori dediğimiz şey; veri, olgular, tipolojiler, taksonomiler veya ampirik bulgular değildir. Olguların bir topluluğu, nasıl bir taş yığını bir ev değilse, bir teori değildir. Benzerce, yapıların bir topluluğu (mesela, yapıların bir tipolojisi) bir teori değildir, çünkü teoriler önermeleri, açıklamaları ve sınırlandırıcı koşulları ihtiva etmek suretiyle yapıların ötesine geçmelidir. Veri, olgular ve bulgular ampirik veya gözlemsel düzeyde işlerler, oysa teoriler kavramsal düzeyde işlerler ve gözlemlerden daha ziyade mantığa dayanırlar.

Araştırmada teorileri kullanmanın pek çok yararı vardır. Birincisi, teoriler, doğal veya sosyal fenomenin meydana gelişinin altında yatan mantığı temin ederler, şu suretle ki hedef fenomenin kilit sürücülerinin (*/key drivers*) ve kilit çıkımlarının neler olduğunu ve hangi alta-yatan proseslerin o fenomeni sürmekten (*/driving*) sorumlu olduğunu ve niçin sorumlu olduklarını açıklayarak. İkincisi, manalandırmaya yardımcı olur, şu suretle ki farklı öğrenme-çalışmalarında iki yapı arasındaki ilişkiye tesir eden olumsal (*/contingent*) faktörlerin keşfedilmesi sayesinde, daha önceki ampirik bulguları teorik bir çerçevede sentezlememize ve çelişkili bulguları uzlaştırmamıza olanak sağlayarak. Üçüncüsü, teoriler gelecek araştırmaya rehberlik eder, şu suretle ki araştırılmaya daha fazla geçecek yapıları ve ilişkileri teşhis etmeye yardımcı olarak. Dördüncüsü, teoriler birikimli bilgi oluşumuna katkıda bulunabilirler, şu suretle ki diğer teoriler arasındaki boşluklar üzerine köprüler kurarak ve mevcut teorilerin yeni bir ışıktaki yeniden değerlendirilmesine neden olarak.

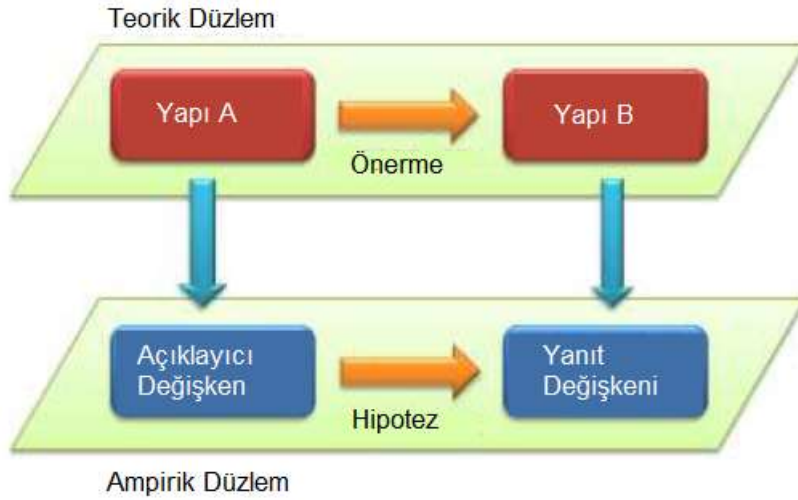
Ancak, teoriler, aynı zamanda, sınırlamalardan paylarına düşeni de alırlar. Gerçeğin basitleştirilmiş açıklamaları olarak, teoriler her zaman ilgilenilen fenomen hakkında kifayetsiz açıklamaları temin edemeyebilirler, çünkü yapıların ve ilişkilerin sınırlandırılmış bir kümesine dayanırlar. Teoriler basit ve parsimonik açıklamalar olarak dizayn edilirken, gerçek anlamıyla (*/significantly*) daha karmaşık olabilir. Dahası, teoriler, araştırmacıların “görüş alanını” perdeleyebilirler veya sınırlandırabilirler, şu suretle ki teori tarafından tanımlanmayan önemli kavramları gözden kaçırmalarına neden olarak.

Bir Teorinin Ana Unsurları

David Whetten (1989) bir teorinin dört ana unsuru haiz olduğunu ifade etmektedir: yapılar, önermeler, mantık ve sınırlandırıcı koşullar/varsayımlar. Yapılar teoriler tarafından “ne” sorusuna verilen cevaptır (yani, bir fenomeni açıklamak için hangi kavramlar önemlidir), önermeler “nasıl” sorusuna verilen cevaptır (yani, bu kavramlar birbirleriyle nasıl ilişkilidir), mantık “niçin” sorusuna verilen cevaptır (yani, bu kavramlar niçin ilişkilidir) ve sınırlandırıcı koşullar/varsayımlar “kim, ne zaman ve nerede” sorularına verilen cevaptır (yani, bu kavramlar ve ilişkiler hangi koşullar altında işler). Yapılar ve önermeler daha önce Bölüm 2’de tartışılmış olsa dahi, tamamlamak adına, burada yeniden betimleyeceğiz.

Yapılar, yüksek bir soyutlama düzeyinde spesifikleştirilen soyut kavramlardır ki ilgilenilen fenomeni açıklamak için spesifik olarak seçilirler. Bölüm 2 den anımsayacağın üzere, yapılar mesela ağırlık ya da yaş gibi tek-boyutlu (yani, tek bir kavramdan ibaret) veya mesela kişilik veya kültür gibi çok-boyutlu (yani, çok sayıda alta-yatan kavramdan oluşan) olabilirler. Mesela yaş, öğrenim ve firma büyüklüğü gibi bazı yapılar kolay anlaşılabilirken, mesela yaratıcılık, önyargı ve organizasyonel çeviklik gibi başka yapılar daha karmaşık ve zor anlaşılabilir olabilirler; ve bunlarla beraber mesela güven, tutum ve öğrenme gibi daha başka yapılar kararlı durumlardan ziyade geçici eğilimleri temsil ediyor olabilirler. Her halükarda, tüm yapılar, açık ve net operasyonel tanımlara sahip olmalıdırlar, öyle ki bu tanımlar yapının nasıl ve hangi analiz düzeyinde (birey, grup, organizasyon ve benzeri) ölçüleceğini tam olarak spesifikleştirmelidir.

Soyut yapıların ölçülebilir temsilleri **değişkenler** olarak adlandırılır. Mesela, zeka düzeyi (IQ skoru) zeka denen soyut bir yapıyı ölçmek maksadıyla teklif edilen bir değişkendir. Daha önce belirtildiği üzere, bilimsel araştırma iki düzlem boyunca ilerler: teorik bir düzlem ve ampirik bir düzlem. Yapılar teorik düzlemde kavramsallaştırılırken, değişkenler ampirik (gözlemsel) düzlemde operasyonelleştirilir ve ölçülür. Dahası, değişkenler Bölüm 2 de tartışıldığı üzere, bağımsız, bağımlı, medyatör veya moderatör olabilirler. Yapılar (teorik düzeyde kavramsallaştırılan) ve değişkenler (ampirik düzeyde ölçülen) arasındaki ayrım Figür 4.1 de gösterilmektedir.



Figür 4.1 Teorik ve ampirik kavramlar arasındaki ayrım.

Önermeler (/propositions), tümdengelimci mantığa dayanılarak, yapılar arasında postüla edilen çağrışımlardır (/associations). Önermeler deklaratif bir formda ifade edilirler ve ideal olarak bir neden-ve-etki ilişkisini göstermelidirler (mesela, eğer *X* meydana gelirse, o zaman *Y* de meydana gelecektir). Nazarı dikkatini celbetmek isterim ki önermeler konjektürel olabilirler, fakat test-edilebilir olmak ZORUNDALARDIR ve eğer ampirik gözlemlerle desteklenmezler ise, reddedilmeleri gerekir. Mamafih, yapılar gibi, önermeler de teorik düzeyde ifade edilirler ve sadece bu yapıların ölçülebilir değişkenleri arasındaki mütakabil ilişkiyi incelemek suretiyle test edilebilirler. Önermelerin ampirik formülasyonuna, değişkenler arasındaki ilişkilerin bir ifadesi olmak üzere, **hipotez** adı verilir. Önermeler (teorik düzeyde formüle edilen) ve hipotezler (ampirik düzeyde test-edilen) arasındaki ayrım Figür 4.1 de gösterilmektedir.

Bir teorinin üçüncü ana unsuru, **mantıktır** ki mantık, bize, postüla edilen önermeleri gerekçelendirmek (/justify) için lazım olan zemini temin eder. Mantık, bir “tutkal” rolü oynar, şöyle ki teorik yapıları birbirine bağlamak suretiyle, bu yapılar arasındaki ilişkilere mana ve münasebet (/relevance) kazandırır. Mantık, aynı zamanda, “açıklama”yı mümkün kılar ki açıklama bir teorinin göbeğinde yer alır. Mantık olmadan, önermeler hususi (/ad hoc), keyfi ve manasız olacaktırlar ve bir teorinin kalbi olan kohezif bir “önermeler sistemine” bağlanamayacaklardır.

Son olarak, tüm teoriler, değerlere, zamana ve uzaya dair **varsayımlar** ve teorinin nerede uygulanabileceğini ve nerede uygulanamayabileceğini buyuran **sınırlandırıcı koşullar** tarafından kısıtlanırlar. Mesela, birçok ekonomik teori, insanın rasyonel (ya da sınırlıca rasyonel) olduğunu varsayar ve insan davranışını anlamının bir yolu olarak, maliyet ve fayda beklentilerine dayanan yarar maksimizasyonunu kullanır. Aksine, politika bilimi teorileri, insanların rasyonel olmaktan daha çok, politik olduklarını ve kendi profesyonel veya kişisel çevrelerinde, kendi güçlerini en üst düzeye çıkaracak ve başkalarını kontrol edecek şekilde kendilerini konumlandırmaya çabaladıklarını varsayar. Altta yatan varsayımlarının mahiyeti

göz önünde alındığında, ekonomik ve politik teoriler doğrudan karşılaştırılabilir değildir ve eğer araştırmacıların ereği, bir organizasyondaki güç sütürütürünü veya onun evrimini anlamaksa, ekonomik teorileri kullanmamalıdır. Benzerce, teoriler, zımnî kültürel varsayımlara (mesela, bireyci mi yoksa kolektif kültürlere mi uygulanacaklar), geçici varsayımlara (mesela, insan davranışının erken mi yoksa daha geç safhalarında mı uygulanacaklar) ve uzaysal varsayımlara (mesela, her yere değil fakat belirli yerlere mi uygulanacaklar) sahip olabilirler. Eğer bir teori münasipçe kullanılacak ya da test edilecekse, o teorinin sınırlarını teşkil eden zımnî varsayımlarının tümü münasipçe anlaşılmalıdır. Ne yazık ki, teorisyenler zımnî varsayımlarını nadiren açıklıkla ifade ederler, ki bu da teorilerin araştırmaların problem durumlarına sıklıkla yanlış uygulanmasına yol açmaktadır.

İyi Bir Teorinin Vasıfları

Teoriler, karmaşık sosyal gerçeğin basitleştirilmiş ve çoğu zaman kısmi açıklamalarıdır. Böyle olunca, iyi açıklamalar veya kötü açıklamalar olabilir ve vargısal olarak, iyi teoriler veya kötü teoriler olabilir. Bir teorinin “iyiliğini” nasıl değerlendirebiliriz? Farklı araştırmacılar tarafından farklı kriterler önerilmiştir; bu önerilerin daha önemli olanları aşağıda sıralanmıştır:

- **Mantıksal tutarlılık:** Teorik yapılar, önermeler, sınırlandırıcı koşullar ve varsayımlar birbiriyle mantıklı olarak tutarlı mı? Bir teorinin bu “ana unsurlarından” bazıları birbiriyle tutarsızsa (mesela, bir teori rasyonelliği varsayar, fakat bazı yapılar rasyonel olmayan kavramları ortaya koyar), o zaman teori kötü bir teordir.
- **Açıklama gücü:** Bir teori gerçeği ne kadar açıklıyor (veya tahmin ediyor)? İyi teoriler, hedef fenomeni, elbette, rakip teorilerden daha iyi açıklar, ki mesela bu, regresyon denklemlerinde, açıklanan varyans değeri (R-kare) ile ölçülür.
- **Yanlışlanabilirlik:** İngiliz filozofu Karl Popper 1940 larda şunu ifade etti ki bir teori, geçerli olmak için, yanlışlanabilir olmak zorundadır. Yanlışlanabilirlik şunu temin eder ki bir teori potansiyel olarak çürütülebilir, eğer ki ampirik veri, araştırmacı tarafından ampirik olarak test edilmeleri olanaklı teorik önermelerle eşleşmez ise. Başka bir deyişle, teoriler ampirik olarak test edilebilir değilse, teori olamazlar. Mesela, “ısının yüksek olduğu bir gün sıcak bir gündür” gibi totolojik ifadeler, ampirik olarak test edilebilir değildir, çünkü sıcak bir gün, zaten, ısının yüksek olduğu bir gün olarak tanımlanır (ve ölçülür) ve dolayısıyla, böylesi ifadeler teorik bir önerme olarak görülebilir. Yanlışlanabilirlik, öte yandan, rakip açıklamaların varlığını gerektirir, ki bu da yapıların kafi derecede ölçülebilir olmasını sağlar ve saire. Mamafih, dikkat etmelisin ki bir teorinin yanlışlanabilir olduğunu söylemekle, bir teorinin yanlışlanması gerektiğini söylemek aynı şey değildir. Eğer bir teori, ampirik kanıta dayalı olarak sahiden yanlışlanır ise, o zaman muhtemelen kötü bir teori ile başlamışsın demektir!
- **Parsimoni:** Parsimoni, bir fenomenin, ne kadar az sayıda değişkenle, ne kadar çok miktarda açıklandığını inceler. On dördüncü yüzyılda yaşamış İngiliz mantıkçı Ockham’lı Peder William’a atfedilen (ve dolayısıyla “Ockham’ın usturası” ya da “Occam’ın usturası” olarak adlandırılan) bu kavram şunu ifade eder ki gözlenen kanıtı yeterince açıklayan rakip açıklamalar arasında en basit olan (yani, en az sayıda değişken kullanan veya en az sayıda varsayımda bulunan) en iyisidir. Kompleks bir sosyal fenomenin açıklaması, daha fazla sayıda yapı eklemek suretiyle, her zaman daha da iyileştirilebilir. Mamafih, böyle bir yaklaşım, gerçeğin “basitleştirilmiş” ve genellenebilir açıklamaları olmaya niyetlenen bir teoriye sahip olma maksadını akamete uğratar. Parsimoni belirli bir teorideki serbestlik derecesi ile alakalıdır. Parsimonik teoriler daha yüksek serbestlik derecelerine sahiptirler ve bu da onların diğer bağlamlara, ortamlara ve popülasyonlara daha kolayca genelleştirilmelerini sağlar.

Teori İnşa Etmede Farklı Yaklaşımlar

Araştırmacılar teorileri nasıl inşa eder? Steinfeld ve Fulk (1990)³ şu dört yaklaşımı önermektedir. İlk yaklaşım, olayların veya davranışların gözlenen biçimlerine dayanarak, teorileri tümevarımcı olarak inşa etmektir. Bu yaklaşım çoğunlukla “grounded teori inşası” olarak adlandırılır, çünkü teori ampirik gözlemler zeminine oturtulur (*/is grounded*). Bu teknik, araştırmacının gözlemsel ve interpretif yeteneklerine büyük ölçüde bağımlıdır ve ortaya çıkacak teori, subjektif ve onaylanabilir olamaz. Dahası, araştırmacı, gözlenen biçimler (*/patterns*) için tutarlı açıklamalar yapamadığı sürece, olayların belli biçimlerini gözlemek bir teori oluşturmaya kafi değildir. Nitel araştırmayı konu eden daha sonraki bir bölümde, grounded teori yaklaşımını tartışacağız.

Teori inşasında ikinci yaklaşım, ilgilenilen fenomenle alakalı farklı tahminci kümelerini ifade etmek için, önceden tanımlanmış bir çerçeveyi kullanmak suretiyle, aşağıdan yukarıya doğru bir kavramsal analiz yürütmektir. Böyle bir çerçeve, basit bir girdi-proses-çıkış çerçevesi olabilir ki bu çerçevede araştırmacı, mesela ilgilenilen fenomen (yani çıktı) ile potansiyel olarak bağlantılı bireysel, organizasyonel ve/veya teknolojik faktörler gibi farklı girdi kategorilerini arayabilir ve o faktörleri hedef fenomene bağlayan altta-yatan prosesi betimleyebilir. Bu, aynı zamanda, tümevarımcı bir yaklaşımdır ki büyük ölçüde araştırmacının tümevarımcı becerilerine dayanır ve interpretasyon, öğrenilmeye-çalışılan fenomen hakkında araştırmacının önsel bilgisi tarafından yanlılaştırılabilir (*/biased*).

Teori inşa etmede üçüncü yaklaşım, yeni bir bağlamı açıklamak için, halihazırda mevcut olan teorileri genişletmek veya modifiye etmektir, mesela şu suretle ki organizasyonel öğrenmeyi açıklamak için, bireysel öğrenme teorilerini genişletmek gibi. Böyle bir genişletme yaparken, eski teorinin belli kavramları, önermeleri ve/veya sınırlandırıcı koşulları korunabilir ve geri kalanı yeni bağlama uyacak şekilde modifiye edilebilir. Bu tümdengelimci yaklaşım, önceki teorisyenler tarafından geliştirilmiş olan sosyal bilim teorilerinin zengin envanterini baz alır ve mevcut teoriler üzerine yeni teorileri inşa ederek, etkili bir yoldan gider.

Dördüncü yaklaşım, halihazırda mevcut olan teorileri, tamamen yeni bağlamlarda, iki bağlam arasındaki sürtüktürel benzerliklerden yararlanmak suretiyle, uygulamaktır. Bu yaklaşım, analogi vasıtasıyla akıl yürütmeye dayanır ve muhtemeldir ki tümdengelimci bir yaklaşım kullanarak teori inşa etmenin en yaratıcı yoludur. Mesela, Markus (1987)⁴ ağ büyümesi için bir kritik kitle teorisi teklif ederken, bir nükleer patlama ve ağların veya ağ tabanlı işletmelerin kontrolsüz büyümesi arasındaki analogik benzerlikleri kullanmıştır. Nükleer bir patlamanın, nükleer patlamayı sürdürebilmesi için, kritik bir radyoaktif madde kütlesi gerektirmesi gibi, Markus, öne sürmüştür ki bir ağ da, büyümesini sürdürmek için kritik bir kullanıcı kitlesi gerektirir ve böyle bir kritik kitlenin yokluğunda, kullanıcılar ağdan ayrılabilirler ve nihayetinde ağın ölümüne neden olurlar.

Sosyal Bilim Teorilerine Misaller

Bu kısımda, farklı sosyal bilim disiplinlerinden birkaç illüstratif teorinin çok kısa bir gözden geçirmesini sunacağız. Bu teoriler, farklı sosyal davranışların tiplerini açıklamaktadır, öyle ki yapıların, önermelerin, sınırlandırıcı koşulların, varsayımların ve alttayatan mantığın bir

³ Steinfeld, C.W. and Fulk, J. (1990). “The Theory Imperative,” in *Organizations and Communications Technology*, J. Fulk and C. W. Steinfeld (eds.), Newbury Park, CA: Sage Publications.

⁴ Markus, M. L. (1987). “Toward a ‘Critical Mass’ Theory of Interactive Media: Universal Access, Interdependence, and Diffusion,” *Communication Research* (14:5), 491-511.

kümesini kullanmak suretiyle. Dikkatini celbetmek isterim ki aşağıdakiler, bu teorilerin sadece basitleştirilmiş birer takdimidir. Her bir teori hakkında daha fazla detay ve derinlik için, okuyuculara, bu teorilerin orijinal kaynaklarına başvurmaları salık verilir.

Vekalet Teorisi (/Agency theory). Vekalet teorisi (aynı zamanda asil-vekil (/principal-agent) teorisi olarak ta adlandırılır), organizasyonel ekonomi literatüründe klasik bir teoridir, başlangıçta Ross (1973)⁵ tarafından, hedefleri birbiriyle uyuşmayan (mesela, bir işveren ve çalışanları arasında, organizasyonel yöneticiler ve hissedarlar arasında, alıcılar ve satıcılar arasında olduğu gibi) iki taraflı ilişkileri açıklamak için önerilmiştir. Vekalet teorisinin amacı, optimal sözleşmeleri ve söz konusu sözleşmelerin, amaç uyumsuzluğunu minimize etmesine yardımcı olabilecek koşulları spesifikleştirmektir. Bu teorinin ana varsayımları şunlardır ki insanlar kendine çıkarıcı, sınırlıca rasyonel ve riskten kaçınan bireylerdir ve bu teori bireysel veya organizasyonel düzeyde uygulanabilir.

Bu teoride yer alan iki taraf asil (/principle) ve vekildir (/agent); asil, kendi adına belirli görevleri yerine getirmesi için vekili istihdam eder. Asilin amacı, verilen görevin hızlı ve etkili bir şekilde tamamlanması iken, vekilin amacı, kendi hızında çalışmak, risklerden kaçınmak ve şirket çıkarlarından çok, kendi çıkarlarının (mesela, kişisel ödeme gibi) peşinde koşmak olabilir. Bunun vargısı, amaç uyumsuzluğudur. Asilin, vekilin davranışını kafi derecede gözlemede veya vekilin yeteneklerini isabetle değerlendirmede gösterdiği yetersizliğin neden olduğu enformasyon asimetrisi problemleri, belki de, bu problemin doğasını oluşturmaktadır. Böylesi bir asimetri, vekalet problemlerine yol açabilir ki vekil verilen görevi yerine getirmek için gerekli olan çabayı göstermeyebilir (*ahlaki risk* problemi) veya işi elde etmek için uzmanlığını ya da becerilerini yanlış sunmuş olabilir ve beklenen düzeyde performans gösteremeyebilir (*ters seçim* problemi). Mesela aylık maaş gibi, davranış-bazlı tipik sözleşmeler bu sorunların üstesinden gelemez. Bundan ötürü, vekalet teorisi, mesela görev tamamlandıktan sonra ödenecek bir komisyon veya ücret gibi çıktı-bazlı sözleşmelerin ya da davranış-bazlı ve çıktı-bazlı sözleşmeleri kombine eden karışık sözleşmelerin kullanılmasını tavsiye eder. Bir çalışan hisse senedi opsiyonu planı, çıktı-bazlı bir sözleşmeye misal iken, çalışan maaşı, davranış-bazlı bir sözleşmeye misaldir. Vekalet teorisi, aynı zamanda, davranış-bazlı sözleşmelerin etki gücünü ıslah etmek için, asillerin istihdam edebileceği araçlar tavsiye eder. Mesela, ahlaki riskin neden olduğu enformasyon asimetrisine karşı mücadele etmek için, denetleme mekanizmalarına yatırım yapmak (mesela denetçileri işe almak gibi), vekilin performansına bağlı olan yenilenebilir sözleşmeler dizayn etmek (performans değerlendirmesi, sözleşmeyi kısmen çıktı-bazlı yapar) veya daha çok programlanabilir ve böylece daha çok gözlenebilir kılmak için, verilen görevin sütrüktürünü ıslah etmek gibi.

Planlı Davranış Teorisi. Ajzen (1991)⁶ tarafından postüla edilen planlı davranış teorisi (PDT), sosyal psikoloji literatüründe genelleştirilmiş bir insan davranışı teorisidir ki bu teori, çok çeşitli bireysel davranışları öğrenmeye-çalışmak için kullanılabilir. Bireysel davranışın bilinçli akılcı seçimi temsil ettiğini ve kognitif düşünmeyle ve sosyal baskılarla şekillendiğini farzeder. Bu teori postüla eder ki davranışlar, kişinin o davranışa dair niyetine dayanır, ki bu da, haliyle, o davranışa, o davranışa dair subjektif norma ve o davranış kontrol altına alma algısına karşı kişinin tutumunun bir fonksiyonudur (Figür 4.2). Tutum, söz konusu edilen davranışın icrasına dair bireyin pozitif veya negatif hislerinin bütünü olarak tanımlanır, ki bu, sözkonusu davranışların farklı vargılarına dair kişinin inançlarının bir toplamı olarak değerlendirilebilir ve o vargıların arzu edilirliliği mesabesinde ağırlıklandırılır. Subjektif norm, bir kişinin, kendisi için

⁵ Ross, S. A. (1973). "The Economic Theory of Agency: The Principal's Problem," *American Economic Review* (63:2), 134-139.

⁶ Ajzen, I. (1991). "The Theory of Planned Behavior," *Organizational Behavior and Human Decision Processes* (50), 179-211.

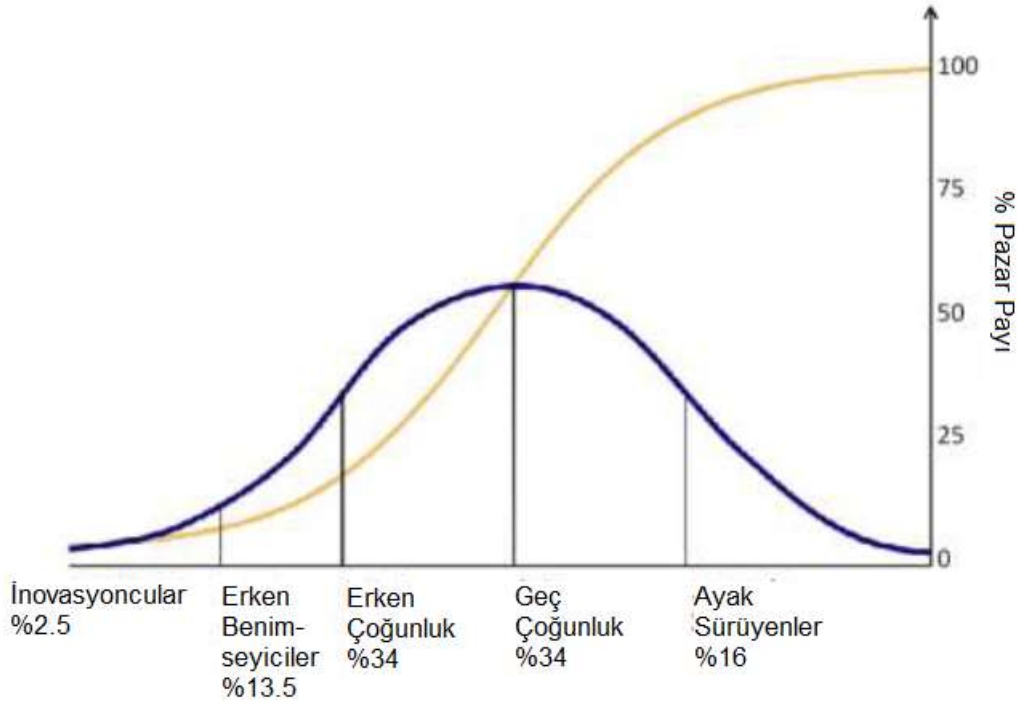
önemli olan insanların, o kişinin söz konusu davranışı icra etmesini bekleyip beklemedikleri hakkındaki algısına tekabül eder ve mesela arkadaşlar, meslektaşlar veya yöneticiler gibi farklı referans gruplarının beklenen normlarının ağırlıklı bir kombinasyonu olarak temsil edilir. Davranışsal kontrol, söz konusu edilen davranışı kısıtlayan içsel veya dışsal kontroller hakkında bir kişinin algısıdır. İçsel kontroller, kişinin söz konusu davranışı icra etme kabiliyetini (öz-etki-gücü) ihtiva edebilirken, dışsal kontrol, bu davranışı icra etmek için gerekli olan dışsal kaynakların elverişliliğine atıf yapar (kolaylaştırıcı koşullar). PDT, aynı zamanda, öne sürer ki insanlar bazen belirli bir davranışı icra etmeye niyetlenebilirler, fakat bunun için ihtiyaç duyulan kaynaklardan mahrum olabilirler; ve bu yüzden, davranışsal kontrol, davranış üzerinde, niyetin müdahil olduğu dolaylı etkisinin yanı sıra, doğrudan bir etkiye sahip olabilir.

PDT, akılcı eylem teorisi olarak adlandırılan daha evvelki bir teorinin bir genişletmesidir ki o teori, niyetin ana sürücülerini olarak tutum ve subjektif normu ihtiva etmekteydi, davranışsal kontrolü değil. Sonki yapı, insanların kendi davranışları üzerinde eksik kontrol sahibi olabildikleri zamanki şartları (mesela web sörfü için yüksek hızda İnternet erişimi olmaması gibi) hesaba katmak için Ajzen tarafından ilave edilmiştir.



Figür 4.2. Planlı davranış teorisi

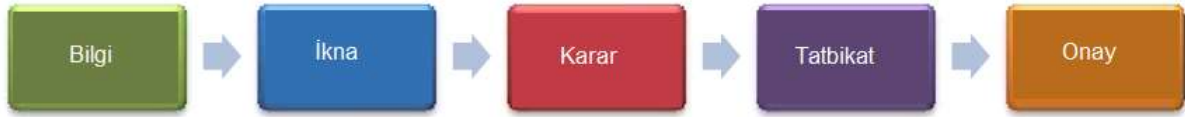
İnovasyon difüzyonu teorisi. İnovasyon difüzyon teorisi (İDT), iletişim literatüründe, seminal bir teodir ki potansiyel benimseyicilerin bir popülasyonunda inovasyonların nasıl benimsendiğini açıklar. Kavramı ilk olarak öğrenmeye-çalışan kişi Fransız sosyolog Gabriel Tarde olmakla beraber, teori, Everett Rogers tarafından 1962 yılında, 508 adet difüzyon öğrenme-çalışmalarına ait gözlemlere dayanılarak, geliştirildi. Bu teoriye dört kilit element şunlardır: inovasyon, iletişim kanalları, zaman ve sosyal sistem. İnovasyonlar yeni teknolojileri, yeni pratikleri veya yeni fikirleri ihtiva edebilir ve benimseyiciler bireyler veya organizasyonlar olabilir. Makro (popülasyon) düzeyde, İDT, inovasyon difüzyonunu, bir iletişim prosesi olarak görür ki bu proseste bir sosyal sistemdeki insanlar, iletişim kanalları (mesela kitlesel iletişim araçları veya önceki benimseyiciler gibi) vasıtasıyla, yeni bir inovasyon ve onun potansiyel faydaları hakkında bilgi sahibi olurlar ve onu benimsemeye ikna edilirler. Difüzyon geçici bir prosedir; difüzyon prosesi birkaç erken benimseyici arasında yavaşça başlar, daha sonra, inovasyon ana akım popülasyon tarafından benimsendikçe hız kazanır ve sonunda benimseyici popülasyonun doygunluğa ulaşmasıyla yavaşlar. Bu nedenle kümülatif benimseme biçimi, Figür 4.3 te gösterildiği üzere S-şekilli bir eğridir ve benimseyici dağılımı bir normal dağılım gösterir. Tüm benimseyiciler birbirinin aynı değildir ve benimseyiciler, benimseme sürelerine göre inovasyoncular, erken benimseyiciler, erken çoğunluk, geç çoğunluk ve ayak sürüyenler olarak tasnif edilebilir. Difüzyon oranı, aynı zamanda, sosyal sistemin karakteristiklerine de bağlıdır, öyle karakteristikler ki mesela kanaat önderlerinin (fikirlerine başkaları tarafından değer verilen uzmanlar) ve değişim ajanlarının (diğerlerinin davranışlarını etkileyen insanlar) mevcudiyeti gibi.



Figür 4.3. S-şekilli difüzyon eğrisi

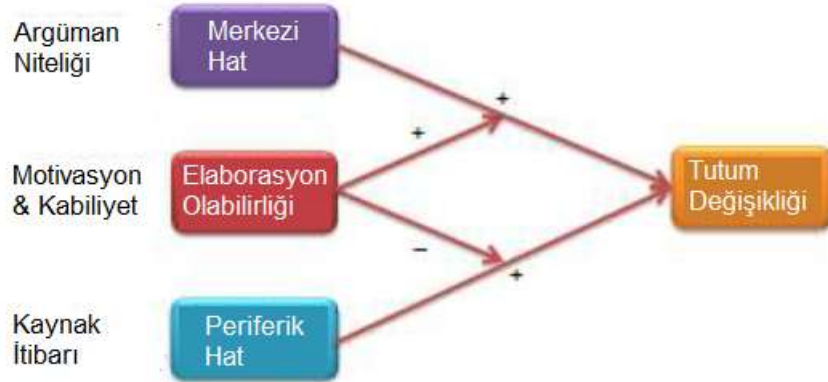
Mikro (benimsiyici) düzeyde, Rogers (1995)⁷, inovasyonun benimsenmesinin beş safhadan oluşan bir süreç olduğunu öne sürmektedir: (1) bilgi: benimsiyicilerin kitle iletişim araçlarından veya kişilerarası kanallardan bir inovasyon hakkında ilk kez bilgi edinmeleri, (2) ikna: önceki benimsiyiciler tarafından inovasyonu denemeye ikna edilmeleri, (3) karar: inovasyonu kabul etmeye veya reddetmeye karar vermeleri, 4) tatbikat: inovasyondan yararlanmaya başlamaları ve (5) onay: en yüksek potansiyeliyle inovasyonu kullanmaya devam etmeye karar vermeleri (Figür 4.4). Benimsiyicilerin inovasyonu benimseme kararlarını şekillendirdiği farzedilen beş inovasyon karakteristiği şunlardır: (1) nispi avantaj: bir inovasyonun önceki inovasyonlara göre beklenen faydaları, (2) uyumluluk: inovasyonun benimsiyicilerin iş görme alışkanlıkları, inançları ve değerleri ile ne derecede uyumlu olduğu, (3) komplekslik: inovasyonun öğrenilmesinin ve kullanılmasının ne decede zor olduğu, (4) denenebilirlik: inovasyonun denemeye dayalı olarak ne derecede test edilebilir olduğu ve (5) gözlenebilirlik: inovasyonu kullanmanın sonuçlarının ne decede açıkça gözlenebilir olduğu. Son iki karakteristik, pek çok inovasyon öğrenme-çalışmasından çoktan beri düşürülmüştür. Komplekslik, inovasyonun benimsenmesi ile negatif korelasyonlu iken, diğer dört faktör pozitif korelasyonludur. İnovasyonun benimsenmesi, aynı zamanda, mesela benimsiyicinin risk alma eğilimi, öğrenim düzeyi, kozmopolitliği ve iletişim tesirliliği gibi kişisel faktörlere de bağlıdır. Erken benimsiyiciler, maceracı, iyi eğitilmiş ve inovasyon hakkında enformasyon edinmek için kitle iletişim araçlarına daha çok itibar edenler iken, geç benimsiyiciler, başlıca enformasyon kaynağı olarak daha çok kişilerarası kaynaklara (mesela arkadaş ve aile gibi) itibar ederler. İDT, “inovasyon-taraftarlığı yanlılığına” sahip olmakla eleştirilmiştir; yani, tüm inovasyonların faydalı olduğunu ve nihayetinde bütün popülasyona difüzyonlarının gerçekleşeceğini (yani yayılacağını) farzetmekle ve mesela hevesler ve modalar gibi etkin olmayan inovasyonların benimsenmeden çabucak yitip gideceğini ya da daha iyi inovasyonların onların yerini alacağını hesaba katmamakla.

⁷ Rogers, E. (1962). *Diffusion of Innovations*. New York: The Free Press. Other editions 1983, 1996, 2005.



Figür 4.4. Inovasyon benimsenme prosesi

Elaborasyon Olabilirlik Modeli. Petty ve Cacioppo (1986)⁸ tarafından geliştirilen, elaborasyon olabilirlik modeli (EOM), psikoloji literatüründe tutum formasyonuna veya değişikliğine dair iki-prosesli bir teoridir. Bireylerin belli bir nesneye, olaylara veya davranışa karşı tutumlarını değiştirmelerine nasıl tesir edilebileceğini ve bu değiştirme stratejilerinin nispi etki gücünü açıklar. EOM farzeder ki bir kişinin tutumu iki tesir “hattınca” şekillendirilebilir, merkezi hat ve periferik hat, ki bu hatlar insanların ihtiyaç duyduğu enformasyon işleme derinliği veya “elaborasyon” bakımından farklılık gösterir (Figür 4.5). Merkezi hat, bir kişinin enformasyonel mesajda yer alan mevzuyla alakalı argümanlar üzerinde düşünmesini ve hedef obje hakkında enforme edilmiş bir yargıya varmadan önce, bu argümanların yararlılıklarını ve münasebetini dikkatle incelemesini gerektirir. Periferik hatta, denekler, hedef nesneye karşı tutumlarını belirlemek için, argümanların niteliğinden daha ziyade, mesela önceki kullanıcıların sayısı, uzmanlardan alınan onaylar veya onaylayıcının beğeniliği gibi dışsal “ipuçlarına” dayanır. Periferik hat, kognitif olarak daha az zorlayıcıdır ve tutum değişikliği hatları, *argüman niteliği* ve *periferik ipuçları* yapılarını sırasıyla kullanmak suretiyle, EOM'de tipik olarak operasyonelleştirilir.



Figür 4.5. Elaborasyon olabilirlik modeli

İnsanların merkezi hatların mı yoksa periferik hatların mı tesiri altında kalacağı, o insanların bir argümanın başlıca yararlılıklarını elabore etme kabiliyetlerine ve motivasyonlarına bağlıdır. Bu elabore etme kabiliyetine ve motivasyonuna “*elaborasyon olabilirliği*” adı verilir. Yüksek elaborasyon olabilirliği durumundaki (yüksek kabiliyet ve yüksek motivasyon) insanlar, sunulan enformasyonu derinlemesine işlemeye daha yatkınlardır ve bu yüzden argüman niteliğinden daha çok etkilenirler, oysa ki düşük elaborasyon olabilirliği durumundakiler periferik ipuçları ile daha fazla motive olurlar. Elaborasyon olabilirliği, durumsal bir özelliktir ve kişisel bir özellik değildir. Mesela, bir doktor tıbbi bir rahatsızlığı teşhis etmek ve tedavi etmek için, merkezi hattı kullanabilir (konuyla ilgili uzmanlığı hasebiyle), ancak otomobilindeki sorunları anlamak için, otomobil tamircilerinden gelen periferik ipuçlarına güvenebilir. Bu itibarla, bu teori, yeni ürünlere veya fikirlere ve hatta sosyal değişime karşı tutum değişikliğinin nasıl gerçekleştirileceğine dair yaygın tatbikata sahiptir.

⁸ Petty, R. E., and Cacioppo, J. T. (1986). *Communication and Persuasion: Central and Peripheral Routes to Attitude Change*. New York: Springer-Verlag.

Genel Caydırıcılık Teorisi. On sekizinci yüzyılın iki faydacı (*/utilitarian*) filozofu Cesare Beccaria ve Jeremy Bentham, Genel Caydırıcılık Teorisi'ni (GCT) hem suçun bir açıklaması olarak ve hem de suçu azaltmak için bir metot olarak formüle ettiler. GCT, niçin belli kişilerin sapkın, anti-sosyal veya kriminal davranışlara angaje olduklarını inceler. Bu teori, insanların esasen rasyonel olduklarını (hem uyumlu hem de sapkın davranışlar için) ve rasyonel bir maliyet-fayda hesaplamasına dayanarak, sapkın davranışları özgürce seçtiklerini savunur. İnsanlar fayda-maksimize-edici davranışları doğal olarak seçtikleri için, kişisel kazanç veya zevk hasıl eden sapkın seçimler, bu tür davranışların cezalandırma formundaki (karşı-önlemler) maliyetlerini artırmanın yanısıra tutuklanma olasılığını da artırmak suretiyle kontrol altına alınabilirler. Cezaların çabukluğu, şiddeti ve kesinliği GCT'de kilit yapılarıdır.

Kriminolojide klasik pozitivist araştırma, mesela yoksulluk, öğrenim eksikliği ve psikolojik koşullar gibi kriminal davranışların genelleştirilmiş nedenlerini bulmaya çalışmaktayken ve suçluları rehabilite etmek için mesela meslek eğitimi vermek ve tıbbi tedavi temin etmek gibi stratejiler teklif etmekteyken, GCT, kriminal karar verme prosesine ve bu prosese tesir eden durumsal faktörlere odaklanmaktadır. Bu yüzden, bir suçlunun kişisel durumu (mesela kişisel değerleri, refahı ve paraya olan ihtiyacı gibi) ve çevresel bağlam (mesela hedefin nasıl korunduğu, polisin ne kadar etkin olduğu ve suçluların ne oranda tutuklanabildiği gibi) bu karar verme prosesinde kilit rol oynarlar. GCT'nin odak noktası suçluların nasıl rehabilite edileceği ve gelecekteki kriminal davranışlardan nasıl kaçındırılacağı değil, fakat kriminal aktivitelerin nasıl daha az çekici kılınacağı ve buna bağlı olarak suçların nasıl önleneceğidir. Bu sonuca ulaşmak için, mesela silindirli kilit takmak ve kendini savunma becerileri edinmek gibi “hedef zorlaştırma”; mesela belli suçlar için şartlı tahliyeyi kaldırmak, “üç tasallut yasası” (üç ofans için, hatta küçük ve hapis cezası gerektirmeseler bile, hapis yatma zorunluluğu) ve idam cezası gibi yasal caydırıcılar; mesela mahalle gözetim programları, uyuşturucu veya çete suçlarına dair spezial görev kuvvetleri ve polis devriyelerini artırmak gibi vasıtaları kullanmak suretiyle tutuklanma şansının artırılması ve mesela “İzinsiz geçişler kovuşturulacaktır” şeklinde görünürlüğü yüksek uyarılar gibi öğretsel programlar suçların önlenmesinde etkilidir. Bu teorinin yalnızca geleneksel suçlar için değil, fakat, aynı zamanda, mesela içerden enformasyon satışı, yazılım korsanlığı ve yasadışı müzik paylaşımı gibi çağdaş beyaz yakalı suçlar için de ilginç tatbikatları vardır.

Bölüm 5

Araştırma Dizaynı

Araştırma dizaynı, bir ampirik araştırma projesinde, veri toplamaya matuf şümüllü bir plandır. Spesifik araştırma sorularına cevap vermeyi veya spesifik hipotezleri test etmeyi hedefleyen ampirik araştırmanın “şeması”dır ki en az üç prosesi spesifikleştirmek zorundadır: (1) veri toplama prosesi, (2) enstrüman geliştirme prosesi ve (3) örnekleme prosesi. Enstrüman geliştirme ve örnekleme prosesleri sonraki iki bölümde betimlenmektedir. Veri toplama prosesi (ki çoğu zaman üstünkörü “araştırma dizaynı” olarak adlandırılır) ise bu bölümde takdim edilmekte ve Bölüm 9-12 de tafsilatlı betimlenmektedir.

Geniş manada, veri toplama metotları kabaca iki kategoride gruplandırılabilir: pozitivist ve interpretif. Mesela laboratuvar deneyleri ve yoklama araştırması gibi **pozitivist metotlar**, teori (veya hipotezlerin) testini hedeflerken, mesela eylem araştırması ve etnografi gibi **interpretif metotlar**, teori inşasına yöneliktir. Pozitivist metotlar araştırmaya tümdengelimci bir yaklaşımı benimser ki bu yaklaşım bir teori ile başlar ve ampirik veri kullanmak suretiyle, teorik postülları test eder. Interpretif metotlar, aksine, tümevarımcı bir yaklaşımı benimser ki bu yaklaşım veri ile başlar ve gözlenen veriden ilgilenilen fenomen hakkında bir teori türetmeye çabalar. Çoğu kere, bu metotlar, nicel ve nitel araştırmayla yanlışça eşleştirilir. Nicel ve nitel metotlar, toplanan verinin tipine (nicel veri sayısal skorları, metrikleri ve benzerlerini haviyken; nitel veri görüşmeleri, gözlemleri ve benzerlerini ihtiva eder) ve kullanılan analiz tekniğine (mesela regresyon gibi nicel teknikler veya kodlama gibi nitel teknikler) matuftur. Pozitivist araştırma, ağırlıklı olarak nicel veri kullanır, fakat nitel veri de kullanabilir, aynı zamanda. Interpretif araştırma, ağırlıklı olarak nitel veriye dayanır, fakat nicel verinin dahil edilmesinden de zaman zaman faydalanabilir. Bazen, nitel ve nicel verinin ortak kullanımı, kompleks bir sosyal fenomene dair, tek başına herhangi bir veri tipi ile elde edilemeyecek, benzersiz bir içgörüyü sahip olmaya yardımcı olabilir ve dolayısıyla, nitel ve nicel veriyi biraraya getiren **karişik-mod dizaynlar** çoğu zaman hayli arzulanabildir.

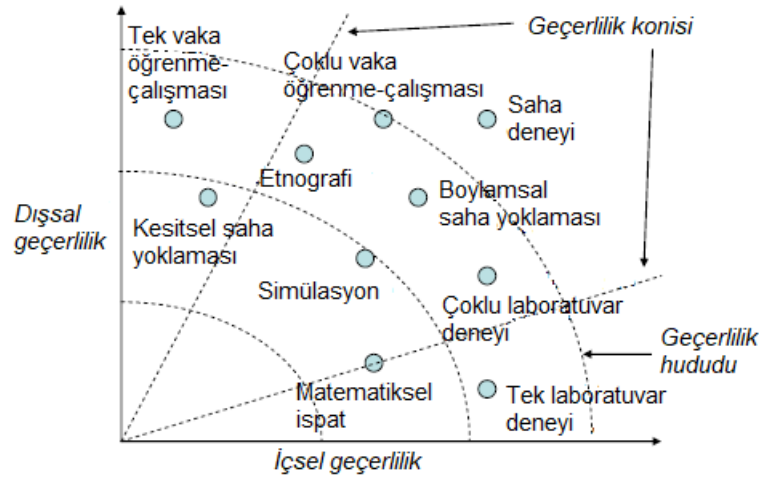
Araştırma Dizaynının Kilit Vasıfları

Araştırma dizaynlarının kalitesi dört kilit dizayn vasfı cinsinden tanımlanabilir: içsel geçerlilik, dışsal geçerlilik, yapı geçerliliği ve istatistiksel netice geçerliliği.

İçsel geçerlilik, aynı zamanda nedensellik denir, bir yanıt değişkeninde gözlenen değişikliğe, araştırma bağlamının dışındaki değişkenlerin değil, gerçekten hipotize edilen açıklayıcı değişkendeki müteakibî değişikliğin neden olup olmadığını inceler. Nedensellik üç koşulun yerine getirilmesini gerektirir: (1) neden ve etkinin kovaryasyonu (yani, neden meydana gelirse, o taktirde etki de meydana gelir; ve eğer neden meydana gelmez ise, etki de meydana gelmez), (2) zamansal öncelik: neden, zaman sürecinde, etkiden önce gelmelidir, (3) inandırıcı görünen alternatif bir açıklama (veya suni korelasyon) olmaması. Mesela laboratuvar deneyleri gibi belli araştırma dizaynları içsel geçerlilik bakımından güçlülerdir, şu sebeple ki açıklayıcı değişken (neden) bir muamele vasıtasıyla manipüle edilebilir ve belli bir zaman noktasından sonra, dışsal değişkenlerin etkileri kontrol altında tutulmak suretiyle, o muamelenin etkisi (yanıt değişkeni) gözlenir. Mesela saha yoklamaları gibi başka dizaynlar içsel geçerlilik bakımından zayıflardır çünkü açıklayıcı değişkeni (neden) manipüle edemezler; ve neden-ve-etki aynı zaman noktasında ölçüldüğü için, zamansal öncelik temin edilemez ve bu şunu eşit derecede olabilir kılar ki beklenen etki beklenen nedene tesir etmiş olabilir, tersi yerine. İçsel geçerliliği diğer

metotlara göre daha yüksek olmasına rağmen, laboratuvar deneyleri, içsel geçerlilik tehditlerine karşı katıyen bağışık değildir ve geçmiş, test etme, enstrümantasyon, regresyon ve daha sonra deney dizaynları bölümünde tartışılacak başka tehditlere karşı hassaslardır. Bununla beraber, farklı araştırma dizaynları kendi nispi içsel geçerlilik düzeyleri bakımından hatırı sayılır değişiklik gösterirler.

Dışsal geçerlilik ya da genelleştirilebilirlik, gözlenen çağrışımların örnekten popülasyona (popülasyon geçerliliği) veya insanlara, organizasyonlara, bağlamlara veya zamana (ekolojik geçerlilik) genelleştirilebilip genelleştirilemeyeceğine matuftur. Mesela, Türkiye'deki finans şirketlerinin bir örneğinden çıkarılan sonuçlar, finans şirketleri popülasyonuna (popülasyon geçerliliği) veya Türkiyedeki diğer şirketlere (ekolojik geçerlilik) genelleştirilebilir mi? Verinin çok çeşitli bireylerden, şirketlerden veya diğer analiz birimlerinden geldiği yoklama araştırması, laboratuvar deneylerine göre daha fazla genelleştirilebilirliğe sahip olma eğilimindedir. Çünkü laboratuvar deneylerinde yapayca tasarlanmış muameleler ve dışsal değişkenler üzerindeki güçlü kontrol, bulguları, muamelelerin ve dışsal değişkenlerin kontrol altına alınabilmediği gerçek yaşam ortamlarına göre daha az genelleştirilebilir hale getirir. Çeşitli araştırma dizaynları için, içsel geçerlilik ve dışsal geçerlilik değişimi Figür 5.1 de gösterilmiştir.



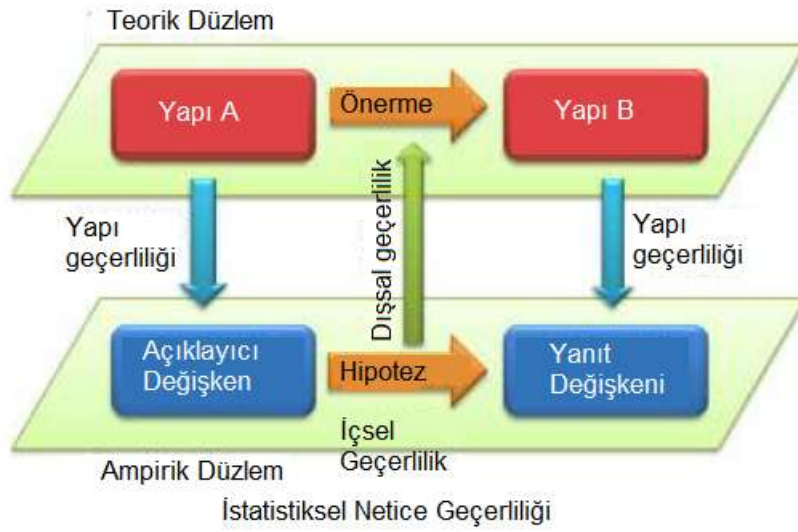
Figür 5.1. İçsel ve dışsal geçerlilik

Bazı araştırmacılar, içsel ve dışsal geçerlilik arasında bir takas olduğunu iddia ederler: daha yüksek dışsal geçerlilik, ancak içsel geçerlilik karşılığında elde edilebilir ve tersi. Ama bu her zaman böyle değildir. Mesela saha deneyleri, boylamsal saha yoklamaları ve çoklu vaka öğrenme-çalışmaları gibi araştırma dizaynları, hem içsel hem de dışsal geçerliliğin daha yüksek derecelerine sahiplerdir. Şahsen, makul düzeyde hem içsel hem de dışsal geçerliliğe sahip, yani Figür 5.1 de gösterilen geçerlilik konisi içinde yer alan araştırma dizaynlarını tercih ederim. Ancak bundan, bu koninin dışındaki dizaynların daha az yararlı veya daha az değerli olduğu yargısı çıkarılmamalıdır. Araştırmacıların dizayn seçimleri, nihayetinde kişisel tercihleri ve yetkinlikleri ve arzu ettikleri içsel ve dışsal geçerlilik düzeyi meselesidir.

Yapı geçerliliği herhangi bir ölçeğin (*measurement scale*) ölçmesi beklenen teorik yapıyı ne kadar iyi ölçtüğünü inceler. Sosyal bilim araştırmalarında kullanılan mesela empati, değişmeye karşı direnç ve organizasyonel öğrenme gibi pek çok yapıyı tanımlamak ölçmekten çok daha zordur. Mesela, yapı geçerliliği bir empati ölçüsünün gerçekten empatiyi ölçtüğünü garantilemelidir, şefkati değil; ki bu zor olabilir çünkü bu yapılar bir şekilde mana itibarıyla

birbirinin benzeridir. Yapı geçerliliği, pozitivist araştırmada, pilot test verisinin korelasyon veya faktör analizine dayanılarak değerlendirilir ki bu bir sonraki bölümde betimlenmektedir.

İstatistiksel netice geçerliliği kullanılan istatistiksel prosedür ile türetilen neticelerin ne derece geçerli olduğunu inceler. Mesela, hipotez testi için doğru istatistiksel metodun kullanılıp kullanılmadığını, kullanılan değişkenlerin o istatistiksel testin varsayımlarını (mesela örnek hacmi veya dağılım gerektirimleri gibi) karşılayıp karşılamadığını vb. inceler. İnterpretif araştırma dizaynları istatistiksel testleri istihdam etmedikleri için, istatistiksel netice geçerliliği böyle bir analize uygulanabilir değildir. Geçerliliğin farklı çeşitleri ve teorik/ampirik düzeylerdeki mevcudiyeti Figür 5.2 de gösterilmektedir.



Figür 5.2 Bilimsel Araştırmada Farklı Geçerlilik Tipleri

İçsel ve Dışsal Geçerliliği Artırmak

En iyi araştırma dizaynları onlardır ki içsel ve dışsal geçerliliğin yüksek düzeylerini garanti edebilirler. Böyle dizaynlar suni korelasyonlara karşı koruma sağlar, hipotez testine olan inancı kuvvetlendirir ve küçük bir örnekten elde edilen sonuçların geniş bir popülasyona genelleştirilebilirliğini garanti eder. Araştırma dizaynlarının içsel geçerliliğini (nedensellik) garantilemek için kontroller gereklidir ve kontrol beş yoldan başarılabılır: (1) manipülasyon, (2) eliminasyon, (3) inklüzyon, (4) istatistiksel kontrol ve (5) rastgeleleştirme.

Manipülasyonda, araştırmacı bağımsız değişkenleri bir veya daha çok düzeyde (“muameleler” adını veririz) manipüle eder ve deneklerin herhangi bir manipülasyona tabii tutulmadığı bir kontrol grubuna karşı muamelelerin etkilerini karşılaştırır. Muameleler yeni bir ilacı veya ilacın farklı bir dozunu, bir öğretim stilini (öğrenciler için) vb. ihtiva edebilir. Bu tip kontrol ancak deneysel veya quasi-deneysel dizaynlar ile başarılabılır, mesela yoklamalar gibi deneysel-olmayan dizaynlar ile değil. Dikkatini çekmek isterim ki eğer denekler farklı düzeylerdeki muamele manipülasyonları arasında kafice ayrımlanamıyor ise deneklerin muamelelere yanıtları farklı olmayabilir ve manipülasyon başarısız olur.

Eliminasyon tekniği dışsal değişkenleri tüm muameleler için sabit tutmak suretiyle elemine etmeye dayanır, mesela öğrenme-çalışmasını tek bir cinsiyetle veya tek bir sosyoekonomik statüyle sınırlandırmak gibi. **İnküzyon** tekniğinde, dışsal değişkenlerin rolü onları araştırma dizaynına dahil etmek ve bağımlı değişken üzerindeki etkilerini ayrıca kestirmek suretiyle

mülâhaza edilir, mesela cinsiyetin (erkek ve dişi) bir faktör olarak yer aldığı faktöriyel dizaynlardaki gibi. Böyle bir teknik yüksek derecede genelleştirilebilirlik temin eder fakat aynı zamanda hatırı sayılır derecede geniş örnekleri gerektirir. **İstatistiksel kontrolde**, dışsal değişkenler ölçülür ve istatistiksel test prosesi esnasında kovaryeteler olarak kullanılır.

Son olarak, **rastgeleleştirme** tekniği ile dışsal değişkenlerin etkilerini bir rastgele örnekleme prosesi marifetiyle götürmek hedeflenir, şu koşulla ki bu etkilerin rastgele bir tabiatı (sistematik-olmayan) olduğu garanti edilebilmelidir. Rastgeleleştirmenin iki tipi şunlardır: (1) **rastgele seçim**, ki bu halde bir örnek bir popülasyondan rastgele seçilir ve (2) **rastgele atama**, ki bu halde popülasyondan rastgele veya rastgele-olmayan bir tarzda seçilen denekler muamele gruplarına rastgele atanır.

Rastgelelik aynı zamanda dışsal geçerliliği de garanti eder, şöyle ki örnekten yapılan çıkarımların örneğin çekilmiş olduğu popülasyona genellenmesini temin eder. Dikkat edilmelidir ki kaynak veya erişim kısıtlarından dolayı rastgele seçim olanaklı olmadığı zaman, rastgele atama zorunludur. Mamafih, popülasyona genellenebilirlikten emin olmak zordur çünkü herhangi bir popülasyon birden çok boyutta farklılık arzedebilir ve biz yalnızca o boyutlardan çok azını kontrol altına alabiliriz.

Popüler Araştırma Dizaynları

Daha evvel ifade edildiği üzere, araştırma dizaynları bilimsel araştırmadaki amaçlarına bağlı olarak iki kategoride sınıflandırılabilir – pozitivist ve interpretivist. Pozitivist dizaynlar teori testine yönelik iken, interpretif dizaynlar teori inşasına yöneliktir. Pozitivist dizaynlar gerçekliğe dair objektif bir bakışa dayanan genelleştirilmiş biçimleri bulmaya çalışırken, subjektif dizaynlar sosyal fenomenlerin subjektif interpretasyonlarını bu fenomeni yaşayan deneklerin perspektifinden bakmak suretiyle bulmaya çalışır. Pozitivist dizaynların bazı popüler misalleri laboratuvar deneylerini, saha deneylerini, saha yoklamalarını, ikincil veri analizini ve vaka araştırmasını haviyken, interpretif dizaynların misalleri vaka araştırması, fenomenoloji ve etnografiyi havidir. Dikkatini celbetmiş olmalı ki vaka araştırması hem teori inşası ve hem de teori testi için kullanılabilir, elbette aynı zamanda değil. Her teknik bilimsel araştırmanın her çeşidi için uygun değildir. Mesela odak gruplar gibi bazı teknikler keşifsel (/exploratory) araştırma için en uygundur, mesela etnografi gibi başka teknikler betimsel (/descriptive) araştırma için en uygundur ve mesela laboratuvar deneyleri gibi daha başkaları ise açıklayıcı (/explanatory) araştırma için idealdir. Müteakiben, bu dizaynlardan bazılarının kısa betimlemesi yapılacaktır. Daha ilave detaylar ise Bölüm 9 ve 12 de temin edilmektedir.

Deneyssel öğrenme-çalışmaları öyledir ki sıkıca kontrol altına alınmış bir ortamda, nedeni etkiden zamanda ayırmak, deneklerin bir grubunu (“muamele grubu”) nedene maruz bırakmak fakat bir diğer grubu (“kontrol grubu”) kendi halinde bırakmak ve bu iki grupta yeralan denekler arasında aritmetik-ortalama etkilerin nasıl değiştiğini gözlemek suretiyle, neden-ve-etki ilişkilerini (hipotezler) test etmeye matuftur. Mesela, eğer belli bir hastalığın tedavisinde, yeni bir ilacın etkililiğini/etkinliğini test etmek için bir laboratuvar deneyi dizayn edeceksek, o hastalığı taşıyan insanların bir rastgele örneğini alabiliriz, bunları iki gruba (muamele ve kontrol grupları) rastgele atayabiliriz, ilacı muamele grubundaki deneklere uygulayabiliriz fakat kontrol grubundakilere sadece plasebo (yani, ilaç değeri olmayan şeker tableti) verebiliriz. Daha kompleks dizaynlar mesela ilacın düşük dozuna karşı yüksek dozu gibi birden çok sayıda muamele grubunu ihtiva edebilirken, mesela ilaç uygulamasının diyetsetel müdahaleler ile biraraya getirilmesi gibi birden çok faktörü de ihtiva edebilir. **Hakiki bir deney dizaynında**, denekler gruplara rastgele atanmalıdır. Eğer rastgele atama yerine getirilmez ise, o taktirde dizayn **yarı-deneyssel (/quasi-experimental)** olur. Deneyler mesela bir üniversitedeki bir

laboratuvar gibi bir yapay ortamda yürütülebilir (laboratuvar deneyleri) veya mesela ilgilenilen fenomenin fiilen cereyan ettiği bir organizasyondaki gibi bir saha ortamında yürütülebilir (saha deneyi). Laboratuvar deneyleri araştırmacının ilgilendiği değişkenleri izole edebilmesini ve dışsal değişkenleri kontrol altına alabilmesini sağlar, ki bu saha deneylerinde olanaklı olamayabilir. Dolayısıyla, laboratuvar deneylerinden yapılan çıkarımlar içsel geçerlilik bakımından daha güçlü olma eğilimi gösterirken, saha deneylerinden yapılan çıkarımlar dışsal geçerlilik bakımından daha güçlü olma eğilimi gösterirler. Deneysel veri nicel istatistiksel teknikler kullanılarak analiz edilir. Deney dizaynının başlıca gücü az sayıdaki değişkeni izole etme, kontrol altına alma ve hassasiyetle inceleme kabiliyeti sayesinde kazandığı güçlü içsel geçerliliği olmakla beraber, başlıca zayıflığı dışsal geçerliliğinin sınırlı olmasıdır ki gerçek hayat çoğu zaman kurmaca laboratuvar ortamlarından daha komplekstir (yani, daha çok sayıda dışsal değişken ihtiva eder). Dahası, eğer araştırma ex ante münasebetli dışsal değişkenleri teşhis edemiyor ve bunları kontrol altına alamıyor ise, böyle bir kontrol eksikliği içsel geçerliliğe de zarar verebilir ve suni korelasyonlara yol açabilir.

Saha yoklamaları deneysel-olmayan dizaynlardır öyle ki bağımsız değişkenleri veya muameleleri kontrol altına almaz veya manipüle etmez, fakat bu değişkenleri ölçer ve istatistiksel metotlar kullanarak etkilerini test eder. Saha yoklamaları, bir yoklama anketi ya da bir sütrüktürlü (/structured) görüşme vasıtasıyla, saha ortamlarındaki deneklerin bir rastgele örneğinden pratiklerin, inançların ya da durumların bir enstantanesini yakalar. **Kesitsel saha yoklamalarında**, bağımsız ve bağımlı değişkenler zamanın aynı noktasında ölçülürken (yani, tek bir anket kullanılarak), **boylamsal saha yoklamalarında**, bağımlı değişkenler bağımsız değişkenlere göre zamanın daha sonraki bir noktasında ölçülür. Saha yoklamalarının güçlü yönleri dışsal geçerlilikleri (veri saha ortamında toplandığı için), çok sayıda değişkeni yakalama ve kontrol altına alma kabiliyetleri ve bir problemi birden çok perspektiften ya da birden çok teori kullanarak öğrenmeye-çalışma kabiliyetleridir. Mamafih, nedeni etkiden zamanda ayırmadıkları için, saha yoklamalarında içsel geçerliliği (neden-ve-etki ilişkileri) tesis etmek zordur ve yoklamalar yanıltıcı yanlışlıklarına (mesela, denekler, kendi yanıtları yerine “sosyal kabul gören” bir yanıt verebilirler) açık olabilirler ki bu da nihayetinde içsel geçerliliğe zarar verir.

İkincil veri analizi öyle bir verinin analizidir ki o veri önceden başka kaynaklar tarafından toplanmıştır ve tablolastırılmıştır. Böyle veri mesela Türkiye İş Kurumunca yayımlanan istihdam istatistikleri veya UNDP’nin ülkeler bazında yayımladığı gelişim istatistikleri gibi devlet kurumlarına ait veriyi, başka araştırmacılar tarafından toplanmış veriyi (çoğu zaman meta-analitik öğrenme-çalışmalarında kullanılır) ya da mesela borsalarda akan finansal veri veya eBay’de akan açıkartırma verisi gibi halka açık üçüncü-parti veriyi ihtiva edebilir. Araştırma için birincil verinin toplanmasının araştırmacının vazifesi olduğu çoğu başka araştırma dizaynlarına bu ters düşer. İkincil veri analizi şu durumda efektif bir araştırma vasıtası olabilir ki birincil veri toplamak çok yüksek maliyetli veya olursuzdur ve ikincil veri araştırmacının sorularını cevaplamak için uygun bir analiz düzeyinde mevcuttur. Bu dizaynın kısıtlılıkları şunlardır ki veri sistematik veya bilimsel bir tarzda toplanmamış olabilir ve bu yüzden bilimsel araştırma için uygun olmayabilir, çünkü söz konusu veri zahiren başka bir maksatla toplanmış olduğundan, araştırmacının sorduğu araştırma sorularına kafi derecede hitap edemeyebilir ve eğer neden ve etki arasında zamansal öncelik bariz değil ise, içsel geçerlilikte problem vardır.

Vaka araştırması bir problemin bir veya birden çok sayıda gerçek hayat ortamında (vaka mekanları) uzun bir zaman periyodu içerisinde derinlemesine soruşturulmasıdır. Veri, görüşmelerin, kişisel gözlemlerin ve içsel veya dışsal dökümanların bir kombinasyonu ile toplanabilir. Vaka öğrenme-çalışmaları tabiatıyla pozitivist (hipotez testi için) veya interpretif (teori inşası için) olabilir. Bu araştırma metodunun güçlü yönü onun ilgilenilen fenomene ilişkin

önceden bilinebilmeyen sosyal, kültürel ve politik potansiyel faktörleri geniş bir yelpazede keşfedebilme kabiliyetidir. Analiz doğal olarak nitele meyleder, ancak ağır bir şekilde bağlamsallaştırılır ve inceliklendirilir. Herhalükarda, bulguların interpretasyonu araştırmacının gözlemsel ve integratif kabiliyetine bağlı olabilir, kontrol eksikliği nedenselliğin tesisini zorlaştırabilir ve tek bir vaka mekanından elde edilen bulgular diğer vaka mekanlarına öylece genelleştirilemeyebilir. Genelleştirilebilirliği artırmak için, bir **çoklu vaka dizaynı** ile başka vaka mekanlarında analiz tekrarlanabilir ve karşılaştırılabilir.

Odak grup araştırması öyle bir araştırma tipidir ki bir lokasyonda küçük bir denek grubu (tipik olarak 6 dan 10 kişiye) teşkil etmeyi ve ilgilenilen bir fenomeni onlara 1.5 ve 2 saat arası bir zaman periyodunda tartıştırmayı ihtiva eder. Tartışma eğitilmiş bir süpervizör tarafından modere edilir ve yürütülür ki bu kişi gündemi belirler ve katılımcıların önüne bir başlangıç soruları kümesi koyar, tüm katılımcıların fikirlerini ve deneyimlerini paylaşabilmelerini temin eder ve katılımcıların yorumları ve deneyimlerine dayalı olarak problem durumuna ait bütünsel bir anlayışı bina etmeye çaba gösterir. Eksik kontrolden dolayı içsel geçerlilik tesis edilemez ve küçük örnek hacminden dolayı bulgular diğer ortamlara genelleştirilemeyebilir. Bu yüzden, odak grupları genellikle açıklayıcı veya betimleyici araştırmada kullanılmazlar, fakat keşifleyici araştırma için çok daha uygunlardır.

Eylem araştırması varsayar ki kompleks sosyal fenomenler, en iyi, bu fenomenlere müdahalelerde ya da “eylemlerde” bulunmak ve bu eylemlerin etkilerini gözlemek suretiyle anlaşılır. Bu metotta, araştırmacı genellikle mesela bir organizasyon gibi bir sosyal bağlamın içine gömülü bir danışman ya da bir organizasyonel üyedir ki mesela karlılığın azalması ya da operasyonel daralmalar gibi bir gerçek problemin çözümüne yönelik mesela yeni organizasyonel prosedürler ya da yeni teknolojiler gibi bir eylemi başlatır. Araştırmacının seçeceği eylemler teoriye dayanmalıdır, öyle ki o eylemlerin arzulanmış değişikliğe niçin ve nasıl neden olabileceğini açıklasın. Sonra, araştırmacı gerektiğinde eylemi modifiye ederek, eylemin sonuçlarını gözlerken, eşzamanlı olarak eylemden öğrenir ve hedef problem ve müdahaleler hakkında teorik içgörüler üretir. Başlangıçsal teori, seçilen eylem hedef problemi ne ölçüde başarıyla çözüyorsa, o ölçüde geçerlenir. Eşzamanlı problem çözme ve içgörü üretimi esas vasıftır ki eylem araştırmasını diğer tüm araştırma metotlarından farklı kılar ve bu yüzden, eylem araştırması araştırma ve pratik arasında köprü kurmak için mükemmel bir metottur. Bu metot hem kendi bağlamının dışında tekrarlanabilmeyen kendine has sosyal problemleri öğrenmeye-çalışmak için de uygundur fakat hem de araştırmacı yanlılığına ve subjektifliğe tabidir ve bulguların genelleştirilebilirliği çoğu zaman öğrenme-çalışmasının yürütüldüğü bağlamla sınırlıdır.

Etnografi interpretif bir araştırma dizaynıdır ki ilhamını antropolojiden almıştır. Şunu önemle vurgular ki araştırma fenomenleri kendi kültür bağlamı içerisinde çalışılmak zorundadır. Araştırmacı uzun bir zaman periyodu boyunca (8 aydan 2 yıla kadar) belli bir kültüre derinlemesine dalmak suretiyle bu periyot içerisinde öğrenilmeye-çalışılan kültürün günlük yaşantısına angaje olur, gözlem yapar ve kayıt tutar. Ve o kültüre ait davranışlar ve o kültürün evrimi hakkında teori kurar. Veri başlıca gözlemsel teknikler, ele alınan kültüre mensup katılımcılar ile formal ve informal interaksiyon ve kişisel saha notları vasıtasıyla toplanırken, veri analizi “deneyimi manalandırmayı” ihtiva eder. Araştırmacı kendi deneyimini öylesine en ince ayrıntısıyla anlatmalıdır ki okuyucu aynı kültürü, orada bulunmasına lüzum kalmaksızın deneyimleyebilsin. Bu yaklaşımın avantajları bağlama duyarlılığı, zengin ve nüanslı bir anlayışı üretmesi ve minimal yanıltıcı yanlılığıdır. Herhalükarda, bu aynı zamanda aşırı zaman ve kaynak yoğun bir yaklaşımdır ve bulgular ele alınan bir kültüre spesifik olduğu için başka kültürlerle genelleştirilebilirliği daha azdır.

Araştırma Dizaynlarının Seçimi

Yukarıdaki araştırma dizaynlarının çokluğuna bakılırsa, araştırmacılar kendi araştırmaları için hangi dizaynı seçmelidir? Genel manada konuşursak, araştırmacılar öyle araştırma dizaynlarını seçmek temayülü gösterirler ki kendilerini en rahat hissetsinler ve yürütmekte en yeterli olsunlar. Fakat, ideal olarak, seçim öğrenilmeye-çalışılan araştırma fenomeninin mahiyetine bağlı olmalıdır. Araştırmanın önsel safhalarında, araştırma probleminin henüz tam netleşmediği ve araştırmacının belli bir araştırma probleminin mahiyetini ve kapsamını tayin etmek istediği zaman, bir odak grup (bireysel analiz ünitesi olarak) veya bir vaka öğrenme-çalışması (organizasyonel analiz ünitesi olarak) keşifsel araştırma için ideal bir stratejidir. Araştırmanın nüfuz alanına daha derinlemesine daldıktan sonra, eğer araştırmacı görürse ki ilgilenilen fenomeni açıklayacak iyi bir teori yoktur ve o alanda mevcut açığı doldurmak için bir teori inşa etmek isterse, mesela vaka araştırması veya etnografi gibi interpretif dizaynlar yararlı dizaynlar olabilir. Eğer rakip teoriler mevcut ise ve araştırmacı bu farklı teorileri test etmek ya da onları daha büyük bir teori içerisinde entegre etmek isterse, mesela deney dizaynı, yoklama araştırması veya ikincil veri analizi gibi pozitivist dizaynlar daha mütenasiptir.

Seçilen spesifik araştırma dizaynı hangisi olursa olsun, araştırmacı nicel ve nitel veri toplamak için çaba göstermelidir öyle ki bu iş için mesela anketler, görüşmeler, gözlemler, dökümanlar ya da ikincil veri gibi tekniklerin bir kombinasyonunu kullanacaktır. Mesela, nicel veri toplamaya matuf bir yüksek derecede sütrüktürlü yoklama anketinde dahi, araştırmacı nitel veri toplamak gayesiyle bir kaç açık uçlu soruya yer verebilir ki sütrüktürlü nicel veri ile asla elde edilemeyecek beklenmedik içgörüler üretebilsin. Benzerce, vaka araştırması ile çoğu nitel veriyi toplamak için çoğu zaman yüz-yüze görüşmeler istihdam edilirken, nicel veri toplama potansiyeli ve değeri gözardı edilmemelidir. Bir misal olarak, organizasyonel karar verme proseslerinin bir öğrenme-çalışmasında, vaka görüşmecisi mesela belli organizasyonel kararları vermenin kaç ay aldığı, anılan karar verme sürecine kaç kişinin dahil olduğu ve kaç alternatif kararın mütala edildiği gibi sayısal nicelikleri kaydedebilir ki bunlar görüşülenlerin anlatısal yanıtlarından elde edilmesi mümkün olmayan değerli içgörüler temin edebilir. Hangi spesifik araştırma dizaynı istihdam edilirse edilsin, araştırmacının ereği mümkün olabildiğince çeşitli veri toplamak olmalıdır ki ilgilenilen fenomen hakkında en iyi olanaklı içgörülerini üretmesine yardımcı olabilsin.

Bölüm 6

Yapıların Ölçümü

Teorik önermeler soyut yapılar arasındaki ilişkilerden meydana gelir. Teorileri (yani, teorik önermeleri) test etme bu yapıları isabetlice, doğru olarak ve bilimsel bir tavırla ölçmeyi gerektirir ki daha sonra aralarındaki ilişkilerin gücü test edilebilsin. Ölçme gerçek dünyanın dikkatli, tasarlanmış gözlemlerine karşılık gelir ve ampirik araştırmanın esasıdır. Sosyal bilim araştırmasında mesela kişinin yaşı, ağırlığı veya firmanın büyüklüğü gibi bazı yapıları ölçmek kolay olabilse dahi, mesela yaratıcılık, önyargı veya yabancılaştırma gibi başka yapıları ölçmek ziyadesiyle zor olabilir. Bu bölümde, böylesi yapıların ölçülerini yaratmak için birbiriyle ilişkili kavramsallaştırma ve operasyonelleştirme proseslerini inceleyeceğiz.

Kavramsallaştırma

Kavramsallaştırma bir zihinsel prodestir ki bulanık ve kesinliksiz yapılar (kavramlar) ve onları oluşturan bileşenler bu proses vasıtasıyla somut ve kesin terimlerle tanımlanır. Mesela, “önyargı” kelimesini sıkça kullanırız ve bu kelime zihnimize belli bir imge uyandırır; mamafih, zorlanabiliriz eğer bizden terimin tam olarak ne manaya geldiğini tanımlamamız istenirse. Eğer birisi başka ırksal gruplar hakkında kötü şeyler söylerse, bu ırksal önyargı mıdır? Eğer aynı vazife için kadınlar adamlardan daha az kazanıyor ise, bu cinsiyet önyargısı mıdır? Eğer kilise müdavimleri inanıyorsa ki inançsızlar cehennemde yanacaktır, bu dinsel önyargı mıdır? Önyargının farklı çeşitleri var mıdır, ve eğer öyle ise, onlar nelerdir? Tüm bu soruları cevaplamak önyargı yapısını doğruca ölçmenin anahtarıdır. Önyargı kavramı içine neyin girdiğini ve neyin girmediğini anlama prosesi işte kavramsallaştırma prosesidir.

Kavramsallaştırma prosesi pek çok sosyal bilim yapılarının kesinsizliği, müphemliği ve muğlaklığı bakımından çok daha fazla önem arzeder. Mesela, “şefkat” “empati” veya “sentimentalite” ile aynı şey midir? Eğer “şefkat empatiyle pozitif ilişkilidir” ifadesine sahip bir önermen varsa, bu önermeyi test edemezsin ta ki empatiyi şefkatten kavramsal olarak ayırabilinceye ve sonra bu iki çok benzer yapıyı ampirik olarak doğruca ölçüncüye kadar. Eğer çok dindar insanlar inanıyorsa ki toplumun mesela inançsızlar, eşcinseller ve kürtaj yapan doktorlar gibi üyeleri günahlarından dolayı cehennemde yanacaklardır, ve cehenneme gitmelerine engel olmak için o “günahkarları” davranışlarını değiştirmeye zorlayan bir uğraşı içine girerlerse, bir önyargısal tavırla mı yoksa şefkat tavrıyla mı davranıyor olurlar? Böylesi yapılara ait tanımlarımız herhangi bir objektif kritere dayanmaz, fakat daha ziyade bu yapılara ait bizim zihinsel imajlarımız (kavramsallaştırmalar) arasındaki müsterek (“inter-subjektif”) bir mutabakata dayanır.

Mesela önyargı veya şefkat gibi yapıları tanımlarken, şunu anlamalıyız ki bazen bu yapılar gerçek değildir veya bağımsızca var olamazlar fakat sadece zihnimizin imgesel kreasyonlarıdır. Mesela, dünyada önyargıdan ari belli kabileler olabilir ve hatta bu kavramın neye tekabül ettiğini zihinlerinde canladırıyor dahi olabilirler. Fakat gerçek hayatta, bu kavrama gerçek olarak muamele etmeye meylederiz. Zihinsel yapıları gerçek olarak addetme prosesine *şeyleştirme* (*reification*) denir, ki yapıları tanımlamada ve onları ölçmek için ölçülebilir değişkenleri tespit etmede merkezi rol oynar.

Yapıları kavramsallaştırmada verilmesi gereken önemli bir karar tekboyutlu mu yoksa çokboyutlu mu olduklarını spesifikleştirmektir. **Tekboyutlu** yapılar öyle yapılardır ki tek bir alttayan boyuta sahip oldukları kabul edilir. Böyle yapılar tek bir ölçü ya da test kullanılarak

ölçülebilirler. Mesela bir kişinin ağırlığı ve rüzgarın hızı gibi basit yapılar ve hatta mesela özgüven (eğer özgüveni tek boyuttan ibaret olarak kavramsallaştırsak, ki bu gerçekdışı bir varsayım olabilir) gibi kompleks yapılar buna misal olarak verilebilir. **Çokboyutlu** yapılar iki veya daha çok sayıda alttayan boyuttan meydana gelir. Mesela, eğer bir kişinin akademik yetkinliğini iki boyut – matematiksel ve sözel yetenek – içerecek halde kavramsallaştırsak, o zaman akademik yetkinlik çokboyutlu bir yapıdır. Bu durumda, diyelim, matematiksel ve sözel yetenek için farklı testler kullanılmak suretiyle alttayan boyutların her biri ayrıca ölçülmelidir ve akademik yetkinlik yapısının genel bir değerini elde etmek için bu iki skor belki ağırlıklandırılmış olarak kombine edilebilir.

Operasyonelleştirme

Teorik bir yapıyı tanımladıktan sonra, onu tam olarak nasıl ölçeriz? **Operasyonelleştirme** bu yapıları ölçmek için indikatörler veya maddeler geliştirme prosesini ifade eder. Mesela sosyoekonomik statü gibi gözlenebilir-olmayan teorik bir yapı, eğer ailenin gelir düzeyi ile tanımlanırsa, yanıtlayıcılara “ailenizin yıllık geliri nedir?” sorusunu sormaktan müteşekkil bir indikatör kullanmak suretiyle operasyonelleştirilebilir. Sosyal bilim yapılarında içkin olan yüksek düzeydeki subjektifliği ve kesinsizliği göz önüne alırsak, bu yapıların çoğunu (mesela yaş, cinsiyet, öğrenim ve gelir gibi bir kaç demografik yapı hariç) ölçmek için çoklu indikatörlerin kullanılması temayülü öne çıkar. Bu proses bize bu indikatörler arasındaki yakınlığı, isabetliliklerinin bir değerlendirmesi (güvenilirlik) olarak, inceleme olanağı sağlar.

Teorik düzeyde kavramsallaştırılan yapıların aksine, indikatörler ampirik düzeyde işler. Belirli bir yapıyı ampirik düzeyde temsil eden indikatörlerin kombinasyonu bir **değişken** olarak adlandırılır. Önceki bir bölümde ifade edildiği üzere, değişkenler, bir araştırma öğrenme-çalışmasında nasıl istihdam edildiklerine bağlı olarak, açıklayıcı değişkenler, yanıt değişkenleri, medyatör değişkenler veya moderatör değişkenler olabilirler. Her indikatör, aynı zamanda, çeşitli **vasıflara** (veya düzeylere) sahip olabilir ve her vasıf bir **değeri** temsil eder. Mesela, “cinsiyet” değişkeni iki vasa sahip olabilir: adam ve kadın. Benzerce, bir müşteri memnuniyeti ölçeği beş vasfı temsil etmek üzere inşa edilebilir: “ziyadesiyle memnuniyetsiz”, “biraz memnuniyetsiz”, “nötr”, “biraz memnun”, “ziyadesiyle memnun”. Vasıfların değeri nicel (sayısal) veya nitel (sayısal-olmayan) olabilir. Nicel veri, mesela regresyon veya sütrüktürel eşitlik modellemesi gibi teknikler kullanılarak analiz edilebilirken, nitel veri mesela kodlama gibi nitel veri analizi tekniklerini gerektirir. Şunu not etmek gerekir ki sosyal bilim araştırmasında pek çok değişken niteldir, hatta bir şekilde sayısal olarak ifade edilmiş olasalar dahi. Mesela, bir müşteri memnuniyet indikatörünü beş vasıflı olarak yaratabiliriz: “ziyadesiyle memnuniyetsiz”, “biraz memnuniyetsiz”, “nötr”, “biraz memnun”, “ziyadesiyle memnun”. Ve bu beş vasıf için sırasıyla 1 den 5 e kadar sayıları atayabiliriz. Böylece nicel veri analizi için sofistike istatistiksel araçları kullanabiliriz. Mamafih, dikkat etmelisin ki sayılar yalnızca yanıtlayıcının kendi memnuniyetine dair kişisel değerlendirmesine ilişkin etiketlerdir ve alttayan değişken (memnuniyet), bu şekilde nicel olarak ifade edilmiş olsa dahi, hala niteldir.

İndikatörler, reflektif veya formatif olabilirler. Bir **reflektif indikatör** öyle bir ölçüdür ki bir alttayan yapıyı “yansıtır”. Mesela, eğer dindarlık bir kişinin ne kadar dindar olduğunu ölçen bir yapı olarak tanımlanırsa, o zaman dinsel faaliyetlere katılım dindarlığın bir reflektif indikatörü olabilir. Bir **formatif indikatör** öyle bir ölçüdür ki bir alttayan yapıyı “teşkil eder” veya ona katkıda bulunur. Böyle indikatörler ilgilenilen yapının farklı boyutlarını temsil edebilirler. Mesela, eğer dindarlık bir inanç boyutundan, bir kulluk boyutundan ve bir ritüel boyutundan müteşekkil olarak tanımlanırsa, o zaman bu farklı boyutları ölçmek için seçilen indikatörler formatif indikaörler olarak telakki edilir. Tekboyutlu yapılar reflektif indikatörler kullanılarak ölçülürken (bununla beraber, özgüven gibi çetrefilli yapıları ölçmek için çoklu

indikatörler kullanılabilir), çokboyutlu yapılar çoklu boyutların bir formatif kombinasyonu olarak ölçülür, her ne kadar alttayatan boyutların her biri bir veya daha çok sayıda reflektif indikatör kullanılarak ölçülebilir olsa dahi.

Ölçüm Düzeyleri

Bir yapıyı operasyonelleştirirken verilecek ilk karar istenilen ölçüm düzeyinin hangisi olacağıdır. **Ölçüm düzeyleri**, aynı zamanda **reyting ölçekleri** (*rating scales*) olarak ta adlandırılır, bir indikatörün alabileceği değerlere atıfta bulunur (fakat indikatörün kendisi hakkında hiç birşey söylemez). Mesela, adam ve kadın (ya da A ve K, ya da 1 ve 2) “cinsiyet” indikatörünün iki düzeyidir. Psikolog Stanley Smith Stevens, *Science* ta 1946 da yayımlanan “On the theory of scales of measurement” adlı seminal makalesinde, bilimsel ölçümler için reyting ölçeklerinin dört jenerik tipini tanımladı: adlandırma, sıralama, aralık ve oran. Bu ölçeklerin istatistiksel özellikleri Tablo 6.1 de görülmektedir.

Ölçek	Merkezi Temayül	İstatistikler	Transformasyonlar
Adlandırma	Mod	Ki-kare	Bire-bir (eşitlik)
Sıralama	Ortanca	Yüzdelik, parametrik-olmayan istatistikler	Monoton artış (sıra)
Aralık	Aritmetik ortalama, gidim, standart sapma	Korelasyon, regresyon varyans analizi	Pozitif doğrusal (afin)
Oran	Geometrik ortalama, harmonik ortalama	Varyasyon katsayısı	Pozitif benzerlikler (çarpımlı, logaritmik)
Not: Daha yüksek mertebeli ölçeklerin tümü daha alçak mertebeli ölçeklere ait istatistikleri kullanabilirler.			

Tablo 6.1 Reyting ölçeklerinin istatistiksel özellikleri.

Adlandırma (/nominal) ölçekleri, aynı zamanda kategorik ölçekler de denir, kategorik veriyi ölçerler. Bu ölçekler karşılıklı engelleyen vasıflara sahip değişkenler veya indikatörler için kullanılır. Misal olarak cinsiyet (iki değer: adam veya kadın), endüstri tipi (fabrika, finans, tarım vb.) ve dinsel mensubiyet (Hristiyan, Müslüman, Yahudi vb.) verilebilir. Mesela adam için 1 ve kadın için 2 gibi her bir değere farklı bir sayı atansa dahi bu sayılar hiç bir manaya gelmez (yani, 1 burada 2 den az veya yarısı değildir) ve mesela adam için M ve kadın için K gibi sayısal-olmayarak ta basitçe temsil edilebilirler. Adlandırma ölçekleri farklı vasıf değerlerine isimler veya etiketler vermekten ibarettir. Adlandırma ölçeği için mütenasip merkezi temayül ölçüsü moddur. Ne aritmetik-ortalama ne de ortanca tanımlanabilir değildir. Uygulanabilecek istatistikler Ki-kare ve frekans dağılımıdır ve yalnızca bire-bir (eşitlik) transformasyon (yani, 1=adam, 2=kadın) yapılabilir.

Sıralama (/ordinal) ölçekleri öyle ölçeklerdir ki sıraya-dizilmiş veriyi ölçerler, mesela bir sınıftaki öğrencilerin mezuniyet not ortalamalarına veya sınav skorlarına göre birinci, ikinci, üçüncü vd. olarak sıralanması gibi. Mamafih, vasıfların mutlak veya nispi değerleri veya vasıf değerleri arasındaki fark değerlendirilemez. Mesela, bir sınıftaki öğrencilerin sıralaması onların mutlak mezuniyet not ortalamaları veya sınav skorları hakkında hiç bir şey söylemez veya birbirlerine göre ne kadar farklı bir performansları olduğunu da söylemez. Fen bilimlerinde klasik bir misal minerallerin sertliğine ait Moh ölçeğidir ki çeşitli minerallerin sertliğini diğer mineralleri çizme kabiliyetlerine göre karakterize eder. Mesela, elmas

yeryüzündeki diğer tüm doğal mineralleri çizebilir ve bu yüzden “en sert” mineraldir. Mamafih, sözkonusu ölçek bu minerallerin fiili sertliğini göstermez hatta onların sertlikleri arasındaki farkların miktarını değerlendirmeye dahi olanak vermez. Sıralama ölçekleri aynı zamanda mesela “kötü”, “orta” ve “iyi” ya da “ziyadesiyle memnuniyetsiz”, “biraz memnuniyetsiz”, “nötr”, “biraz memnun” ve “ziyadesiyle memnun” gibi vasıf etiketlerini (çapalar) kullanabilir. İkincisinde, “biraz memnun” yanıtlayıcıların “ziyadesiyle memnun” yanıtlayıcılardan daha az memnun olduğunu söyleyebiliriz, fakat memnuniyet düzeylerini niceleyebilmeyiz. Bir sıralama ölçeğinin merkezi temayülü onun ortancası veya modu olabilir, ancak aritmetik-ortalamanın interpretasyonu yapılamaz. Bu yüzden, istatistiksel analiz yüzdelikler ve parametrik-olmayan teknikleri ihtiva edebilir, fakat mesela korelasyon, regresyon ve varyans analizi gibi sofistike teknikler mütenasip değildir. Monoton artan transformasyonlar (ki sıralamayı korur) uygulanabilir.

Aralık (/interval) ölçekleri öyle ölçeklerdir ki bunlarda ölçülen değerler yalnızca sıralı değil, fakat aynı zamanda bitişik vasıflar için eşit mesafelilerdir. Mesela, sıcaklık ölçeği Fahrenheit veya Celcius olabilir ki her biri için 30 ve 40 derece arasındaki fark 80 ve 90 derece arasındaki fark ile aynıdır. Benzerce, eğer bir ölçek yanıtlayıcıların yıllık gelirlerini şu vasıfları (aralıkları) kullanarak ölçüyor ise: 0 TL den 10000 TL ye, 10000 TL den 20000 TL ye, 20000 TL den 30000 TL ye ilh., bu da yine bir aralık ölçeğidir, çünkü her aralığın orta noktası (yani, 5000 TL, 15000 TL, 25000 TL vb.) birbirlerinden eşit uzaklıklardadır. Zeka puanı (IQ) hakeza bir aralık ölçeğidir, çünkü ölçek öyle dizayn edilmiştir ki 100 ve 110 luk IQ skorları arasındaki farkın 110 ve 120 arasındaki ile aynı olması beklenir (bunun ne kadar doğru olduğunu gerçi bilmesek te). Aralık ölçeği bize bir vasıfın diğerine göre “ne kadar daha fazla” olduğunu inceleyebilmeyi temin eder, ki bu durum adlandırma ve sıralama ölçekleri ile olanaklı değildir. Merkezi temayül ölçüsü olarak aritmetik-ortalama, ortanca veya mod kullanılabilir. Yayılımı ölçmek için mesela gidim ve standart sapma uygundur. Mütenasip istatistiksel analizler adlandırma ve sıralama ölçekleri için mütenasip olanların tümüne ilaveten korelasyon, regresyon, varyans analizi ve diğerlerini ihtiva eder. Mütenasip ölçek transformasyonları pozitif linear olanlardır. Dikkat etmelisin ki daha önce sözkonusu olan memnuniyet ölçeği katıyyen bir aralık ölçeği değildir, çünkü “ziyadesiyle memnun” ve “biraz memnun” arasındaki farkın “nötr” ve “biraz memnun” arasındaki farka veya “biraz memnuniyetsiz” ve “ziyadesiyle memnuniyetsiz” arasındaki farka eşit olup olmadığını söyleyebilmeyiz. Mamafih, sosyal bilim araştırmacıları çoğu zaman bu farklar eşitmiş gibi davranırlar (doğru değil elbette) ki sıralama ölçeği verisini analiz etmek için istatistiksel teknikleri kullanabilsinler.

Oran (/ratio) ölçekleri öyle ölçeklerdir ki adlandırma, sıralama ve aralık ölçeklerinin tüm niteliklerine sahip olmalarına ilaveten aynı zamanda “bir “gerçek sıfır” noktasına da sahiplerdir (ki burada sıfır değeri alttayatan yapının yokluğunu veya mevcudiyetsizliğini ifade eder). Fen bilimlerinde ve mühendislikte mesela kütle, bir düzlemin eğimi ve elektrik yükü gibi pek çok ölçüm oran ölçeğine tabiyken, mesela bir organizasyona mensubiyet süresi ve firma hacmi (çalışan sayısı veya toplam gelir olarak ölçülebilir) gibi bazı sosyal bilim değişkenleri oran ölçeğine tabidir. Mesela hacmi sıfır olan bir firmanın hiç çalışanı olmadığı veya hiç geliri olmadığı anlaşılır. Kelvin sıcaklık ölçeği, Fahrenheit ve Celcius’un aksine yine bir oran ölçeğidir, çünkü bu ölçeğin sıfır noktası (-273.15 derece Celcius’a eşittir) keyfi bir değer olmaktan ziyade maddenin partiküllerinin kinetik enerjisinin sıfır olduğu sıcaklığa tekabül eder. Bu ölçeklerin “oran” ölçeği olarak adlandırılmalarının sebebi iki ölçüm noktasının oranının manalı olması ve interpretasyonunun yapılabilir olmasıdır. Mesela, 10 işçilik hacmi olan bir firma 5 işçi hacmindeki bir firmanın iki katıdır, bunu 10000 kişiyi istihdam eden bir firma için 5000 kişi istihdam eden başka bir firmaya nispeten aynen söyleyebiliriz. Tüm merkezi temayül ölçüleri, geometrik ve harmonik aritmetik-ortalamalar da dahil, oran ölçekleri ile kullanılabilir. Yine mesela standartlaştırılmış gidim (/studentized range) veya varyasyon katsayısı gibi tüm

yayılm ölçüleri kullanılabilir. Tüm istatistiksel metotlar hakeza kullanılabilir. Mesela çarpımlı veya logaritmik transformasyonlar gibi pozitif transformasyonlar yine kullanılabilir.

Yukarıda tartışılan dört jenerik ölçek tipini temel alarak, sosyal bilim araştırmaları için spesifik reyting ölçekleri (*rating scales*) geliştirebiliriz. Yaygın istihdam edilen reyting ölçekleri arasında ikili (*binary*), Likert, semantik diferansiyel veya Guttman ölçeklerini sayabiliriz. Daha başka az yaygın ölçekleri burada anmadık.

İkili (*binary*) ölçekler. İkili ölçekler adlandırma ölçekleridir ki mesela evet veya hayır, var veya yok, doğru veya yanlış vb. gibi iki değerden birini alabilen ikili maddelerden müteşekkildir. Mesela, “politik aktivizm” yapısı için tipik bir ikili ölçek altı ayrı ikili maddeden müteşekkil olabilir (Tablo 6.2 ye bak). Bu ölçekteki her madde bir ikili maddedir ve bir yanıtlayıcıya ait “evet”lerin toplam sayısı (0 dan 6 ya kadar bir değer) o kişinin politik aktivizminin bütüncül bir ölçüsü olarak kullanılabilir. Bu maddelerin nasıl türetildiğini anlamak için, bu bölümde ileride yer alan “Ölçeklendirme” kesimine bakmalısın. İkili ölçekler mesela cinsiyet için erkek veya dişi, istihdam için tam-zamanlı veya kısmi-zamanlı vb. gibi başka değerleri de alabilir. Eğer bir istihdam durumu maddesi ikiden fazla olanaklı değeri almak üzere (mesela, işsiz, tam-zamanlı, kısmi-zamanlı ve emekli) modifiye edilirse, artık o ikili değildir fakat hala bir adlandırma ölçeği maddesidir.

Bir kamu görevlisine hiç mektup yazdın mı?	Evet	Hayır
Bir politik bildiriye hiç imzaladın mı?	Evet	Hayır
Bir politik dava için hiç para bağışladın mı?	Evet	Hayır
Seçimlerde adaylığını koyan bir politikacı için hiç para bağışladın mı?	Evet	Hayır
Bir gazete veya dergi editörüne hiç bir politik mektup yazdın mı?	Evet	Hayır
Bir kişiyi oy verme kararını değiştirmesi için hiç ikna ettin mi?	Evet	Hayır

Tablo 6.2 Politik aktivizmi ölçmeye matuf altı maddeli ikili bir ölçek.

Likert ölçeği. Rensis Likert (1932) tarafından dizayn edildi. Sosyal bilim araştırmalarında sıralama verisinin ölçümünde çok popüler bir reyting ölçeğidir. Bu ölçek basitçe ifade edilmiş Likert maddelerini ihtiva eder ki yanıtlayıcılar bu ifadelere ne derecede katıldıklarını veya katılmadıklarını “kesinlikle katılmıyorum”dan “kesinlikle katılıyorum”a doğru giden beş veya yedi-noktalı bir ölçek üzerinde belirtebilirler. “Çalışan özgüveni” için altı-maddeli tipik bir Likert ölçeği yapısı Tablo 6.3 te misallendirilmiştir. Likert ölçekleri toplamalı ölçeklerdir, yani genel ölçek skoru yanıtlayıcıların her bir madde için seçtikleri ölçek noktalarına atfedilen değerlerin bir toplamı olabilir.

Likert ölçeği maddeleri, ikili maddelere göre daha fazla ayrıntılılığa (daha incelikli yanıt) izin verir çünkü yanıtlayıcıların ifadeye karşı nötr/tarafsız olması halini de ihtiva eder. Üç veya dokuz değer (ki “çapa” denir bunlara çoğu zaman) kullanılabilir, fakat değerlerin sayısının tek olması bir “nötr” (yani, “ne katılmıyorum ne de katılıyorum”) çapanın mevcudiyetini sağlar. Bazı öğrenme-çalışmalarında nötr orta-puanı çıkartıp, değerlerin sayısını çift yaparak, bir “zorlanmış seçim yaklaşımı” kullanılmıştır ki bunun gayesi yanıtlayıcıları Likert ifadesine katılmamaya veya katılmaya zorlamaktır. Fakat bu iyi bir strateji değildir çünkü bazı insanlar verilen ifadeye karşı gerçekten nötr olabilirler ve zorlanmış seçim yaklaşımı onların nötr duruşunu kaydetme fırsatını kaybettirir. Likert ölçeğinin kilit bir karakteristiği ifadeler farklı maddelere veya indikatörlere göre değişse bile, çapaların (“kesinlikle katılmıyorum”dan

“kesinlikle katılıyorum”a kadar) aynı kalmasıdır. Likert ölçekleri sıralama ölçekleridir çünkü çapaların eşit aralıklı olmaları zorunluluğu yoktur, bazen biz onlara aralık ölçekleri gibi muamele ediyor olsak bile.

	Kesinlikle Katılmıyorum	Kısmen Katılmıyorum	Nötr	Kısmen Katılıyorum	Kesinlikle Katılıyorum
İşimden memnunum	1	2	3	4	5
İşyerindekilerle iyi geçiniyorum	1	2	3	4	5
İşyerinde amirimle ilişkim beni gurulandırıyor	1	2	3	4	5
İşyerindekilerin benimle çalışmaktan memnun olduklarını söyleyebilirim	1	2	3	4	5
İş arkadaşlarımdan bana saygı duyduklarını söyleyebilirim	1	2	3	4	5
İşyerinde yararlı olduğumu düşünüyorum	1	2	3	4	5

Tablo 6.3 Çalışan özgüvenini ölçmek için altı-maddeli bir Likert ölçeği.

Semantik diferansiyel ölçek. Bu bir kompozit (çok-maddeli) ölçektir ki bu ölçekte yanıtlayıcılardan tek bir ifadeye karşı kanaatlerini/görüşlerini veya hislerini/duygularını zıt kutuplar olarak yerleştirilen farklı sıfat çiftlerini kullanarak belirtmeleri istenir. Mesela, “Ulusal sağlık sigortasına karşı tutum” yapısı Tablo 6.4 te gösterilen dört madde vasıtasıyla ölçülebilir.

Likert ölçeğinde olduğu üzere, genel ölçek skoru münferit maddelerin skorlarının bir toplamı olabilir. Şuna dikkatini çekmek istiyorum ki Likert ölçeklerinde ifadeler değişir fakat çapalar her bir maddede aynı kalır. Halbuki, semantik diferansiyel ölçekte ifade sabit kalırken, çapalar (sıfat çiftleri) her bir maddede değişir. İnsanların nesnelere, olaylara veya davranışlara karşı tutumlarını ölçmek için semantik diferansiyelin mükemmel bir teknik olduğuna inanılır.

Ulusal sağlık sigortası hakkındaki kanaatlerinizi nasıl reytinglendirirsiniz?						
	Çok	Biraz	Hiçbiri	Biraz	Çok	
İyi	☐	☐	☐	☐	☐	Kötü
Yararlı	☐	☐	☐	☐	☐	Yararsız
Özenli	☐	☐	☐	☐	☐	Özensiz
İlginç	☐	☐	☐	☐	☐	Sıkıcı

Tablo 6.4 Ulusal sağlık sigortasına karşı tutumu ölçmek için bir semantik diferansiyel ölçek

Guttman ölçeği. Louis Guttman tarafından dizayn edilen bu kompozit ölçek ilgilenilen yapının yoğunluğuna göre artan bir sırada (en az yoğunundan en çok yoğununa doğru) tertip edilen bir maddeler dizisini kullanır. Bir misal olarak, “göçmenlere karşı tutum” yapısı Tablo 6.5 te gösterilen beş madde vasıtasıyla ölçülebilir. Bu Guttman ölçeğindeki her bir madde o maddenin yoğunluğuna göre değişen bir ağırlığa (burada belirtilmeyen) sahiptir ve her bir yanıtın ağırlıklandırılmış kombinasyonu bir gözlemin toplaişik ölçüsü olarak kullanılır.

Aşağıda göçmenler hakkındaki ifadelerle dair kanaatlerini nasıl reytinglendirirsin?		
Göçmelere ülkende vatandaşlık verilmesine karşı mısın?	Evet	Hayır
Göçmenlerin mahallende yaşamasına karşı mısın?	Evet	Hayır
Bir göçmenle kapı komşusu olmaya karşı mısın?	Evet	Hayır
Bir göçmenin yakın arkadaşın olmasına karşı mısın?	Evet	Hayır
Ailenden birinin bir göçmenle evlenmesine karşı mısın?	Evet	Hayır

Tablo 6.5 Göçmenlere karşı tutumu ölçmek için beş-maddeli bir Guttman ölçeği

Ölçeklendirme

Bir önceki alt-bölümde yanıtlayıcıların altta-yatan bir yapıya ait önceden-dizayn-edilmiş maddelere veya indikatörlere yanıtlarının nasıl ölçüleceğini tartıştık. Fakat indikatörlerin kendilerini nasıl yaratıyoruz? İndikatörlerin yaratılması prosesine ölçeklendirme denir. Daha formal olarak, **ölçeklendirme** ölçmenin (ölçüm teorisi) bir dalıdır ki gözlenemeyen yapılar hakkındaki nitel yargıları nicel, ölçülebilir metrik birimlerle irtibatlandırmak suretiyle ölçülerin inşasını ihtiva eder. Stevens (1946) şöyle demişti “Ölçeklendirme nesnelerin bir kurala göre sayılara atanmasıdır.” Bu soyut kavramları somut terimlerle ölçme prosesi ampirik sosyal bilim araştırmasındaki en zor işlerden birisi olarak varlığını devam ettirmektedir.

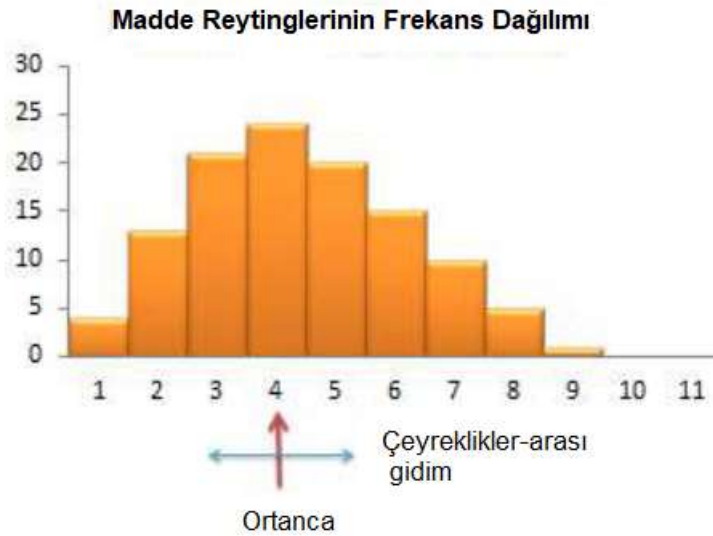
Bir ölçeklendirme prosesinin çıkımı bir **ölçektir**, ki bu belirli bir yapıya ait maddeleri veya indikatörleri ölçmek için ampirik bir yapıdır. Şunu anlamak önem arz ediyor ki bu alt-bölümde tartışılan “ölçekler” bir önceki alt-bölümde tartışılan “reyting ölçeklerinden” biraz farklıdır. Bir reyting ölçeği yanıtlayıcıların belirli bir maddeye karşı reaksiyonlarını kaydetmek için kullanılır, mesela adlandırma düzeyinde ölçeklendirilmiş bir madde evet/hayır reaksiyonunu kaydeder ve sıralama düzeyinde ölçeklendirilmiş bir madde “kesinlikle katılmıyorum” ve “kesinlikle katılıyorum” arasında bir değeri kaydeder. Bir ifadeye veya bir enstrümana bir reyting ölçeği izafe etmek ölçeklendirme değildir. Daha ziyade, ölçeklendirme o maddelere reyting ölçeklerinin izafe edilmesinden önce, ölçek maddelerinin formal olarak geliştirilmesi prosesidir.

Ölçekler tek-boyutlu veya çok-boyutlu olabilir ki bu hal altta-yatan yapının tek-boyutlu (mesela, ağırlık, rüzgar hızı, firma hacmi) veya çok-boyutlu (mesela, akademik yatkınlık, zeka) olmasına bağlıdır. Tek-boyutlu ölçekler yapıları düşükten yükseğe doğru giden tek bir ölçek ile ölçer. Şunu da ifade etmek gerekir ki bu ölçeklerin bazıları birden çok sayıda madde ihtiva edebilir, fakat bu maddelerin tümü aynı altta-yatan boyutu ölçmek içindir. Bu, bilhassa, özgüven gibi düşükten yükseğe doğru giden tek bir boyuta sahip olduğu varsayılan pek çok sosyal bilim yapılarına ait bir durumdur. Çok-boyutlu ölçekler, öte taraftan, ilgilenilen yapının her bir boyutunu ayrı ayrı ölçmek için farklı maddeler veya testler ihtiva eder ve daha sonra ilgilenilen çok-boyutlu yapının genel bir ölçüsünü elde etmek için her boyuta ait skorlar birleştirilir. Mesela, akademik yatkınlık öğrencinin matematiksel ve sözel yetenekleri için ayrı ayrı testler kullanılarak ölçülebilir ve bu skorlar birleştirilerek akademik yatkınlığın genel bir ölçüsü elde edilebilir. Sosyal bilim araştırmasında kullanılan ölçeklerin çoğu tek-boyutlu olduğu için, bu alt-bölümün devamında tek-boyutlu ölçeklerin yaratılmasına dair üç yaklaşımı inceleyeceğiz.

Tek-boyutlu ölçeklendirme metotları yirminci yüzyılın ilk yarısında geliştirildi ve yaratıcılarının isimleriyle anılırlar. En popüler üç tek-boyutlu ölçeklendirme metodu şunlardır:

(1) Thurstone eşit-görünümlü ölçeklendirme, (2) Likert toplamli ölçeklendirme ve (3) Guttman kümülatif ölçeklendirme. Bu üç yaklaşım pek çok bakımdan benzer olmakla beraber ölçek maddelerinin yargıcılar tarafından reytinglendirilmesi ve nihai maddelerin seçiminde kullanılan istatistiksel metotlar bakımında anahtar farklılıkları haizlerdir. Bu metotların her biri burada sırasıyla tartışılacak.

Thurstone eşit-görünümlü ölçeklendirme metodu. Louis Thurstone en ilk ve en ünlü ölçeklendirme teorisyenlerinden biri olarak 1925 yılında eşit-görünümlü aralıklar metodunu yayımladı. Bu metot ilgilenilen yapının açık bir kavramsal tanımı ile başlar. Bu yapıyı ölçmek için, bu tanıma dayanarak, potansiyel ölçek maddeleri üretilir. Bu maddeler ölçülecek yapı hakkında bilgi sahibi olan uzmanlar tarafından üretilir. Aday maddelerin başlangıç havuzu (ideal olarak 80'den 100'e kadar madde) benzer bir tarzda kelimelere dökülmelidir, mesela yanıtlayıcıların katılmayabileceği veya katılabileceği ifadeler olarak (ve sorular veya başka şeyler olarak değil) dile getirmek gibi. Sonra, bu aday maddeler havuzundan ilgilenilen yapıyı temsil edecek spesifik maddeleri seçmek için bir yargıcılar paneli oluşturulur. Yargıcılar enstrüman inşası prosesinde eğitilmiş akademisyenleri veya ilgilenilen yanıtlayıcıların (yani, fenomene aşina olan insanların) bir rastgele örneğini havi olabilir. Seçim prosesi her bir yargıcıya bağımsız olarak her bir maddeyi o maddenin kastedilen yapıyı ne kadar yakından yansıttığına dayanarak 1 den 11 e kadar giden bir ölçekte reytinglendirterek yapılır (1 aşırı aleyhte ve 11 aşırı lehte olmayı temsil eder). Her madde için, ortanca ve çeyreklikler-arası gidim (75^{inci} ve 25^{inci} yüzdelikler arasındaki fark – bir yayılım ölçüsü) hesaplanır, ki bunlar Figür 6.1 de gösterildiği gibi bir histogram üzerine çizilir. Nihai ölçek maddeleri ortancaların bir gidimi boyunca eşit aralıklardaki ifadeler olarak seçilirler. Bunu yapabilmek için ortak bir ortancaya sahip maddeler gruplandırıldıktan sonra, her ortanca grubu içerisinde çeyreklikler-arası gidimi en küçük olan madde seçilir. Mamafih, madde seçimi için tamamen istatistiksel analize güvenmek yerine, aday maddeleri her düzeyde incelemek ve en açık ve en akla yatkın ifadeyi seçmek belki daha iyi bir strateji olabilir. Her bir ölçek maddesinin ortanca değeri bir ağırlığı temsil eder ki bu ağırlık maddelerin ilgilenilen yapıyı temsil edecek kompozit ölçek skoruna katkısıdır. Şimdi bir ölçek sahibiyiz öyle ki bu ölçek üzerindeki 11 noktanın her birine bir madde veya ifadenin tekabül ettiği (ve böyle ağırlıklandırılan) bir cetvel gibi görünür. Maddeler ölçeğin 11 noktası boyunca eşitçe göründükleri için, bu teknik eşit-görünümlü ölçek olarak adlandırılır.



Figür 6.1 Thurstone ölçeği maddelerinin histogramı

Thurstone aynı zamanda tek-boyutlu ölçeklerin inşası için iki ilave metot daha yarattı – *ardışık aralıklar metodu* ve *eşli karşılaştırmalar metodu* – ki her ikisi de eşit-görünümlü aralıklar

metoduna çok benzerdir, farkları yargıcıların veriyi nasıl reytinglendirdikleridir. Mesela, eşli karşılaştırmalar metodu her yargıcının her bir ifade çifti arasında bir yargıda bulunmasını gerektirir (her bir ifadeyi bağımsız olarak 1 den 11 e kadar giden bir ölçekte reytinglendirmek yerine). Zaten adı bu yüzden eşli karşılaştırmalar metodu. Çok sayıda ifade ile, bu yaklaşım, eşit-görünümlü aralıklar metoduyla kıyaslandığında, aşırı zaman alıcı ve kullanışsız olabilir.

Likert toplamı ölçeklendirme metodu. Likert metodu Murphy ve Likert (1939) tarafından geliştirilen tek-boyutlu bir ölçeklendirme metodudur. Bu bölümde betimlenen üç ölçeklendirme yaklaşımının hayli olanaklı ki en popüler olanıdır. Thurstone metodunda olduğu üzere, Likert metodu da ilgilenilen yapının açık bir tanımı ile başlar ve 80 den 100 e kadar giden sayıda potansiyel ölçek maddesi üretmek için bir uzmanlar kümesi istihdam eder. Bu maddeler daha sonra yargıcılar tarafından 1 den 5 e (veya 1 den 7 ye) kadar giden bir reyting ölçeğinde şöylece reytinglendirilir: kavrama kesinlikle katılmıyor ise 1, kavrama kısmen katılmıyor ise 2, kararsız ise 3, kavrama kısmen katılıyor ise 4 ve kavrama kesinlikle katılıyor ise 5. Bu reytinglendirmeyi takiben, nihai ölçekte yer alacak spesifik maddeler şu bir kaç yoldan biri ile seçilebilir: (1) Her bir maddenin yargıcı reytingleri ve maddeler toplamı (her bir yanıtlayıcı için maddelerin tümünü toplayarak elde edilir) arasındaki iki-değişikli korelasyonları hesaplamak suretiyle, madde ve maddeler-toplamı korelasyonu düşük (mesela, 0.60 tan daha az) çıkan maddeleri atmak. (2) Her bir madde için yargıcıların üst çeyreğinin ve alt çeyreğinin aritmetik-ortalamaları arasındaki fark için bir *t*-testi yapmak suretiyle, *t*-değeri büyük çıkan maddeleri (yani, üst ve alt çeyrekteki yanıtlar arasında en iyi ayrımı yapanları) seçmek. Son merhalede, araştırmacı kendi yargısına göre, hem madde ve maddeler-toplamı korelasyonu yüksek hem de ayrımı yüksek (yani, büyük *t*-değerlerine sahip) maddelerin nispeten küçük bir kümesini (diyelim 10 dan 15 kadar sayıda) elde edebilir. Likert metodu tüm maddelerin eşit ağırlıklı olduğunu varsayar, bu sayede, yanıtlayıcının her bir maddeye verdiği yanıtlar o yanıtlayıcı için kompozit bir skor yaratmak için toplanabilir. Zaten, bu metodun adı toplamı ölçektir. Dikkat etmelisin ki yapının orijinal yönüne ters mana taşıyan herhangi bir madde, toplama işleminden önce, ters yönde kodlanmalıdır (yani, 1 olmalı 5, 2 olmalı 4 ve benzeri).

Guttman kümülatif ölçeklendirme metodu. Guttman (1950) tarafından dizayn edilen kümülatif ölçeklendirme metodu Emory Bogardus'un sosyal mesafe tekniğine dayanır. Sosyal mesafe tekniği insanların diğer insanlarla sosyal ilişkiler kurma arzusunun bir yoğunluk derecesine sahip olduğunu varsayar. Guttman kümülatif ölçeklendirme metodu, “en az yoğun”dan “en çok yoğun”a doğru sıralanmış edilmiş maddelerin bir listesini kullanarak, bu yoğunluğu ölçer. Fikir şu ki bu listedeki maddelerin herhangi birine katılan bir kişi aynı zamanda bu maddeden önceki tüm maddelere katılıyor demektir. Pratikte, bu kümülatif biçime mükemmelce uyan bir maddeler listesi çok nadir bulunur. Bir maddeler listesinin kümülatiflik fikrine ne kadar iyi tekabül ettiğini görmek için skalogram analizi yapılır.

Önceki ölçeklendirme metodları gibi, Guttman metodu da ilgilenilen yapının açık bir tanımı ile başlar ve sonra aday maddelerin geniş bir kümesini geliştirmek için uzmanlar istihdam edilir. Sonra, bir yargıcılar grubu aday maddeleri eğer maddeyi yapıya faydalı telakki ederlerse “evet” ve eğer maddeyi faydasız telakki ederlerse “hayır” olarak reytinglendirirler. Sonra, bir matris veya tablo oluşturulur ki tüm aday maddelere yargıcıların yanıtlarını gösterir. Bu matris daha çok “evet” sahibi olan yargıcılar üste ve daha az “evet” sahibi olanlar alta gelecek şekilde yukarıdan aşağıya doğru azalan sırada sıralanır. Aynı sayıda “evet” sahibi olan yargıcılar için, ifadeler soldan sağa doğru, en çok katılımdan en aza göre, sıralanabilir. Tablo 6.6 da böyle bir matrisi görebilirsin. Dikkatini çekmek istiyorum ki ölçek şimdi neredeyse kümülatif, eğer soldan sağa doğru (maddeler boyunca) okunursa. Mamafih, bir kaç istisnada bulunulabilir, Tablo 6.6 da görüldüğü gibi, bu yüzden ölçek tamamen kümülatif değil. Kümülatiflik özelliğine en iyi yaklaşan maddelerin bir kümesini belirlemek için, skologram adı verilen bir veri analizi tekniği kullanılabilir (ya da bu görsel olarak yapılabilir, eğer maddelerin sayısı az ise). Bu

istatistiksel teknik aynı zamanda her bir madde için bir skor kestirimi yapar ki bu skor maddelerin tümünün kümesinde yanıtlayıcının genel skorunu hesaplamada kullanılabilir.

Yanıtlayıcı	Madde 12	Madde 5	Madde 3	Madde 22	Madde 8	Madde 7
29	Y	Y	Y	Y	Y	Y
7	Y	Y	Y	-	<u>Y</u>	-
15	Y	Y	Y	Y	-	-
3	Y	Y	Y	Y	-	-
32	Y	Y	Y	-	-	-
4	Y	Y	-	<u>Y</u>	-	-
5	Y	Y	-	-	-	-
23	Y	Y	-	-	-	-
11	Y	-	-	<u>Y</u>	-	-
<u>Y</u> bu matrisin mükemmelse kümülatif olmasını engelleyen istisnaları belirtiyor						

Tablo 6.6 Bir Guttman ölçeği için sıralanmış reyting matrisi

Endeksler

Bir endeks bir kompozit skordur, öyle ki *çoklu yapıların* (bileşenler adı verilir) ölçülerinin bir kurallar ve formüller kümesi kullanılarak toplaştırılmasından türetilir. Endeksler ölçeklerden şu bakımdan farklıdır ki ölçekler de ölçüleri toplaştırırlar fakat bu ölçüler farklı boyutları veya *tek bir yapının* aynı boyutunu ölçer. Endeks için iyi bilinen bir misal tüketici fiyat endeksidir (CPI) ki bu endeks “the Bureau of Labor Statistics of the U.S. Department of Labor” tarafından her ay hesaplanmaktadır [Türkiye İstatistik Kurumu tarafından hesaplanan TÜFE muadili]. CPI tüketicilerin mal ve hizmetlere genel olarak ne kadar ödemek zorunda olduklarını ölçer ve başlıca sekiz kategoriye bölünmüştür (gıda ve içecek, barınma, eşya, taşıma, sağlık, eğlence, eğitim ve iletişim ve diğer mal ve hizmetler) ki detayında 200 den çok alt maddeye bölünür. Her ay, devlet memurları 80 binden fazla maddenin cari fiyatlarını almak için tüm ülkeyi yoklarlar. Her bir maddenin lokasyonunu ve satın alınma olasılığını hesaba katan karmaşık bir ağırlıklandırma şeması kullanılarak, bu fiyatlar analizciler tarafından kombine edilerek, bir formüller ve kurallar serisi vasıtasıyla genel endeks skorunun kombinasyonuna dahil edilirler.

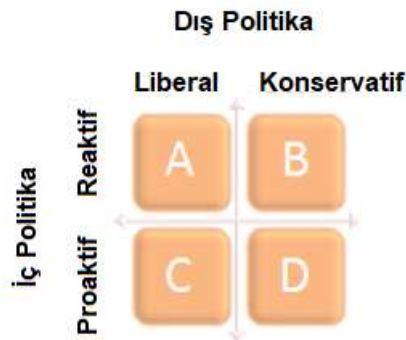
Endekslere başka bir misal sosyo-ekonomik statü endeksidir (SES) ki aynı zamanda Duncan sosyoekonomik endeks (SEI) olarak adlandırılır. Bu endeks üç yapının bir kombinasyonudur: gelir, eğitim ve meslek. Gelir dolar cinsinden ve eğitim yıl veya alınan derece cinsinden ölçülürken, meslek kategorilere veya statü düzeylerine göre sınıflandırılır. Bu birbirinden çok farklı ölçüler, “mesleki eğitim”in (o meslekteki insanların bir veya daha çok yıllık üniversite eğitimi almış olanlarının yüzdesi) ve “mesleki gelir”in (o meslekteki insanların spesifik bir yıllık gelirden daha fazlasını kazananlarının yüzdesi) ağırlıklandırılmış bir kombinasyonu kullanılarak, genel SES endeksi skorunu elde etmek için kombine edilir. Mamafih, SES endeksi ölçümü araştırmacılar arasında pek çok görüş-ayrılığına ve anlaşmazlığa yol açmıştır.

Bir endeks yaratma prosesi bir ölçeğinki ile benzerdir. İlk olarak, endeks kavramsallaştırılır (tanımlanır) ve endeksin bileşenleri oluşturulur. Bu basit görünse bile, endekse hangi bileşenlerin (yapıların) dahil edileceği veya hangilerinin dahil edilmeyeceği hakkında yargıcılar arasında pek çok anlaşmazlık çıkabilir. Mesela, SES endeksinde, gelir eğitim ve meslek ile korelasyonlu değil midir? Eğer öyle ise, bu bileşenlerden yalnızca biri mi dahil edilmelidir yoksa her üçü de mi? Literatürü gözden-geçirmek, teorileri kullanmak ve/veya uzmanlarla veya kilit paydaşlarla görüşmek bu meseleyi halletmeye yardımcı olabilir. İkinci olarak, her bir bileşen operasyonelleştirilir ve ölçülür. Mesela, meslekler nasıl kategorize edilecek, bilhassa çünkü zaman içerisinde bazı mesleklerde farklılıklar olabilir (mesela, İnternet’ten önce Web geliştiriciler yoktu)? Üçüncü olarak, Endeks skorunu hesaplamak için bir kural veya formül yaratılır. Yine, bu proses fazlaca subjektiflik barındırabilir. Son olarak, mevcut veri veya yeni veri kullanılarak endeks skoru geçerlenir.

Endeksler ve ölçekler, her ne kadar, ilgilenilen yapıyı temsil edecek tek bir sayısal skor veya değer üretiyor olsalar dahi, bunlar pek çok bakımdan birbirlerinden farklıdır. Birincisi, endeksler çoğu zaman birbirinden çok farklı olan (mesela, SES endeksindeki gelir, eğitim ve meslek) ve farklı yollardan ölçülen bileşenlerden meydana gelirler. Mamafih, ölçekler tipik olarak benzer maddelerin bir kümesini ihtiva ederler ki bunlar için aynı reyting ölçeği kullanılır (mesela, beş-noktalı Likert ölçeği gibi). İkincisi, endeksler çoğu zaman mesela fiyatlar veya gelir gibi objektif olarak ölçülebilen değerlerin bir kombinasyonuyken, ölçekler mesela tutum, önyargı veya özgüven gibi subjektif veya yargısal yapıları değerlendirmek için dizayn edilirler. Bazıları ölçeklendirme metodolojisinin sofistikasyonunun ölçekleri endekslerden ayırdığını iddia ederken, başka bazıları endeksleme metodolojisinin de eşit derecede sofistikeleştirilebileceğini idda etmektedir. Herhalükarda, hem endeksler hem de ölçekler sosyal bilim araştırmasında esansiyel araçlardır.

Tipolojiler

Ölçekler ve endeksler tek-boyutlu yapıların sıralama ölçülerini üretirler. Mamafih, araştırmacılar bazen, kategorilerin veya tiplerin bir kümesini yaratmak için, iki veya daha çok sayıdaki yapıya ait ölçüleri özetlemeyi arzu ederler. Buna **tipoloji** adı verilir. Ölçekler veya endekslerin aksine, tipolojiler çok-boyutlulardır fakat yalnızca adlandırılmış değişkenleri ihtiva ederler. Mesela, gazetelerin politik bir tipolojisini yaratmak için, Figür 6.2 de gösterildiği üzere, gazetelerin editoryal köşelerinde ifade ettikleri üzerinden iç ve dış politikaya karşı oryantasyonlarına dayanılabilir. Bu tipoloji gazeteleri dört “ideal tip”ten (A dan D ye) birine kategorize etmek için, gazetelerin bu ideal tiplere dağılımını belirlemek için ve hatta belki gazeteleri başka vasıflarına bağlı olarak bu dört ideal tipten birine sınıflandıracak bir sınıflandırma modeli yaratmak için kullanılabilir.



Figür 6.2 Gazetelerin çok-boyutlu bir tipolojisi

Özet

Konuyu bitirirken, řunu söyleyelim ki sosyal bilimlerde ölçek inřası çeřitli kilit kararları ihtiva eden kompleks bir prosestir. Bu kararların bazıları řunlardır:

- Bir ölçek mi, endeks mi yoksa tipoloji mi kullanmalısın?
- Veriyi nasıl analiz etmeyi planlıyorsun?
- Arzuladıđın ölçüm düzeyi (adlandırma, sıralama, aralık veya oran) veya reyting ölçeđi ne?
- Kaç adet ölçek noktası kullanmalısın (mesela, 1 den 10 a; -3 ten +3 e)?
- Ölçek noktalarının sayısı tek mi yoksa çift mi olmalı (yani, bir nötr veya orta-nokta değeri istiyor musun)?
- Ölçek noktalarını nasıl etiketlemek istersin (bilhassa semantik diferansiyel ölçekler için)?
- Nihayetinde, ölçek maddelerini veya endeks bileřenlerini üretmek için hangi prosedürü (mesela, Thurstone, Likert veya Guttman) kullanabilirsin?

Bölüm 7

Ölçek Güvenilirliği ve Geçerliliği

Önceki bölümde, sosyal bilim araştırmasında yapıları ölçmenin birtakım zorluklarını inceledik. Mesela, “empatiyi” değil de “şefkat”i ölçüyor olduğumuzu nasıl bileceğiz? Çünkü, her iki yapı da mana olarak az çok benzerdir. Ya da şefkat empati ile aynı şey midir? Durumu daha kompleks hale getiren bu yapıların bazen imgesel kavramlar olmaları (yani, gerçekte var olmamaları) ve çok-boyutlu olmalarıdır (ki bu halde onları oluşturan boyutların teşhis edilmesi problemi ile de karşı karşıya kalırız). Böylece, sosyal bilim yapılarını, tercih ettiğimiz herhangi bir ölçeği kullanarak öylece ölçmek kafi değildir. Aynı zamanda, kullanacağımız ölçeklerin şunları sağlayıp sağlamadığını test etmeliyiz: (1) Ölçek ölçmek istediğimiz gözlenebilir yapıyı mı sahiden ölçüyor? (Yani, ölçek geçerli mi?) (2) Ölçek kastedilen yapıyı tutarlıca ve kesince ölçüyor mu? (Yani, ölçek güvenilir mi?) Güvenilirlik ve geçerlilik birer mihenk taşıdır ki bilimsel araştırmada ölçüm prosedürlerinin kifayetleri ve isabetlilikleri bunlarla değerlendirilir. Bu ikisine müştereken bir ölçeğin “psikometrik özellikleri” adı verilir.

Bir ölçü güvenilir olabilir fakat geçerli olmayabilir, eğer ki bir şeyi çok tutarlıca ölçüyor fakat yanlış yapıyı tutarlıca ölçüyorsa. Aynı şekilde, bir ölçü geçerli olabilir fakat güvenilir olmayabilir, eğer ki doğru yapıyı ölçüyor fakat tutarlıca ölçemiyor ise. Figür 7.1 de gösterildiği üzere, bir hedefe atış yapma analojisini kullanırsak, bir yapının hem güvenilir hem de geçerli çok-maddeli bir ölçüsü hedef üzerinde merkezileşmiş ve sıkıca öbekleşmiş atışlardan müteşekkil olacaktır. Geçerli fakat güvenilir olmayan bir ölçü hedef üzerinde merkezileşmiş fakat sıkıca öbekleşmemiş, daha ziyade hedefin etrafına saçılmış atışlardan müteşekkil olacaktır. Son olarak, güvenilir fakat geçerli olmayan bir ölçü, sıkıca öbekleşmiş fakat hedefin uzağındaki atışlardan müteşekkil olacaktır. Böylece, ilgilenilen yapıların kifayetli ölçümlerini garanti etmede hem güvenilirliğe hem de geçerliliğe ihtiyaç vardır.



Figür 7.1 Güvenilirliğin ve geçerliliğin karşılaştırılması

Güvenilirlik

Güvenilirlik (*/reliability*) bir yapının ölçüsünün tutarlılığının (*/consistency*) ya da itimat edilebilirliğinin (*/dependability*) derecesidir. Başka bir deyişle, eğer aynı yapıyı çok defa ölçmek için bu ölçeği kullanırsak, altta-yatan fenomenin değişmediğini varsayarak, her defasında hemen hemen aynı sonucu elde eder miyiz? Güvenilmez (*/unreliable*) ölçümün bir misali insanların senin ağırlığını sanılamasıdır. Hayli muhtemel ki, insanların sanıları farklı olacaktır, farklı ölçüler tutarsız olacaktır ve bu yüzden de, “sanılama” ölçüm tekniği güvenilmez olacaktır. Daha güvenilir bir ölçüm bir ağırlık ölçeği (yani bir tartı) kullanmak olabilir ki bu

ölçeğin (yani tartının) üzerine her çıktığında aynı değeri elde edebilirsin, meğerki senin ağırlığın ölçümler arasında hakikaten değişiyor olmasın.

Dikkatini çekmek istiyorum ki güvenilirlik tutarlılığı (/consistency) kasteder, isabetliliği (/accuracy) değil. Bir önceki ağırlık ölçeği misalinde, eğer ağırlık ölçeği yanlış kalibre edilirse (diyelim, daha iyi hissedesin diye, senin ağırlığından beş kilogram eksiltiyor olsun), senin hakiki ağırlığını ölçemeyecek ve bu yüzden geçerli bir ölçü olmayacaktır. Öyle de olsa, yanlış kalibre edilmiş ağırlık ölçeği sana her defasında hala aynı ağırlığı verecektir (ki bu da senin hakiki ağırlığının beş kilogram eksisidir) ve böylece ölçek güvenilirdir.

Sosyal bilim ölçümlerinde güvenilmez gözlemlerin kaynağı nedir? Başlıca kaynaklardan birisi gözlemcinin (ya da araştırmacının) subjektifliğidir. Eğer bir firmada çalışan morali, çalışanların birbirlerine gülümseyip gülümsemedikleri, birbirleriyle şakalaşp şakalaşmadıkları ve benzeri davranışları izlenmek suretiyle ölçülüyor ise, çalışanları çok yoğun bir günde (şakalaşacak veya laflayacak zamanları yok iken) veya az yoğun bir günde (daha şen-şakrak ve daha konuşkan iken) izleyen farklı gözlemciler farklı moral ölçüleri çıkarabilirler. İki gözlemci, yine, aynı günde farklı moral düzeyleri çıkarabilirler, neyi şaka olarak gördüklerine ve neyi şaka olarak görmediklerine bağlı olarak. “Gözlem” nitel bir ölçüm tekniğidir. Bazen, nicel ölçümler kullanılmak suretiyle, güvenilirlik artırılabilir, mesela, (ters yönde) bir moral ölçüsü olarak bir ay boyunca yapılan şikayetleri saymak gibi. Elbette, şikayetlerin sayısı geçerli bir moral ölçüsü olabilir veya olmayabilir, fakat insan subjektifliğine daha az tabidir ve bu sebeple daha güvenilirdir. Güvenilmez gözlemin ikinci bir kaynağı kesinliksiz (/imprecise) ve muğlak sorular sormaktır. Mesela, eğer insanlara gelirlerinin ne olduğunu sorarsan, farklı yanıtlayıcılar bu sorunun interpretasyonunu aylık gelir, yıllık gelir veya saatlik ücret olarak farklı yapabilirler, ve böylece, ortaya çıkacak gözlemler hayli ayrı ve güvenilmez olabilecektir. Güvenilmezliğin üçüncü bir kaynağı yanıtlayıcıların pek aşına olmadıkları veya pek umursamadıkları meseleler hakkında sorular sormaktır. Mesela, bir Türk üniversite öğrencisine Kanada’nın Slovenya ile ilişkisinden memnun olup olmadığını sormak veya bir şirketin personel müdüründen o şirketin teknoloji stratejisinin etkililiğini reytinglendirmesini istemek gibi.

Öyle ise, güvenilir ölçüleri nasıl yaratabiliriz? Eğer bizim ölçümümüz başkalarından enformasyon talep etmeyi ihtiva ediyorsa, ki çoğu sosyal bilim araştırmasında durum budur, o zaman araştırmacı subjektifliğine daha çok bağlı olan veri toplama teknikleri (mesela gözlemler gibi) yerine subjektifliğe daha az bağlı olanları (mesela anket gibi) tercih ederek, yanıtlayıcıların cevabını bilebilecekleri veya umursadıkları meselelere dair sorular sorarak, ölçülerimizde muğlak maddelerden kaçınarak (mesela, yıllık geliri öğrenmek isteyip istemediğimizi açıkça ifade ederek) ve indikatörlerimizin lafzını basitleştirmek suretiyle bazı yanıtlayıcılar tarafından yanlış anlaşılmalarmını sağlayarak (mesela, manaları bilinebilmeyecek zor kelimelerden kaçınarak) başlayabiliriz. Bu stratejiler ölçülerimizin güvenilirliğini artırabilir, öyle ki ölçümlerin tamamen güvenilir olmasını temin edemeyecek olsalar bile. Ölçüm enstrümanlarının güvenilirliği, her halükarda, test edilmelidir. Güvenilirliği kestirmenin pek çok yolu vardır ki şimdi onları tartışacağız.

Reytingciler-arası güvenilirlik. Reytingciler-arası güvenilirlik aynı zamanda gözlemciler-arası güvenilirlik diye adlandırılır ki aynı yapıyı reytinglendiren iki veya daha çok sayıda bağımsız reytingci (gözlemci) arasındaki tutarlılığın bir ölçüsüdür. Çoğu zaman, bu bir plot öğrenme-çalışması ile değerlendirilir ve ilgilenilen yapının ölçüm düzeyine bağlı olarak iki yoldan yapılabilir. Eğer ölçü kategorik ise, tüm kategorilerin bir kümesi tanımlanır, reytingciler her bir gözlemin hangi kategoriye düştüğünü kaydederler ve reytingciler arasındaki uyuma yüzdesi reytingciler-arası güvenilirliğin bir kestirimi olur. Mesela, 100 gözlemi olanaklı üç kategori içerisinde reytinglendiren iki reytingci varsa ve bunların reytingleri gözlemlerin %75 inde eşleşiyor ise, o zaman reytingciler-arası güvenilirlik 0.75 tir. Eğer ölçü aralık veya oran

ölçeğinde ise (mesela, sınıf içi aktivite iki reytingçi tarafından her 5 dakikada bir 1 den 7 ye kadarlık bir yanıt ölçeği ile ölçülüyor), o zaman iki reytingçinin ölçüleri arasında bir basit korelasyon yine reytingçiler-arası güvenilirliğin bir kestirimi olarak hizmet eder.

Test-yenidentest güvenilirliği. Test-yenidentest güvenilirliği aynı yapının aynı örneğe iki farklı zaman noktasında uygulanan iki ölçümü (testi) arasındaki tutarlılığın bir ölçüsüdür. Eğer gözlemler iki test arasında esaslı farklılık göstermiyorsa, o halde ölçü güvenilirdir. İki test arasındaki gözlemlerin korelasyonu test-yenidentest güvenilirliğinin bir kestirimidir. Burada şuna dikkat etmelisin ki iki test arasındaki zaman aralığı kritiktir. Genel olarak, geçen süre ne kadar uzun olursa, iki gözlemin bu süre zarfında farklılık göstermesi şansı o kadar yükselebilir ve test-yeniden test güvenilirliği o kadar azalır.

İki-yarı güvenilirliği. İki-yarı güvenilirliği bir yapı ölçüsünün iki yarısı arasındaki tutarlılığın bir ölçüsüdür. Mesela, eğer elimizde belirli bir yapının on-maddeli bir ölçüsü varsa, bu on maddeyi rastgele beşerli iki kümeye böleriz (eğer total madde sayısı tek ise, iki yarım eşit sayıda olmayabilir) ve enstrümanı bir bütün olarak yanıtlayıcıların bir örneğine uyguluruz. Daha sonra, her bir yanıtlayıcı için her bir yarının total skorunu ve her bir yarıdaki total skorlar arasındaki korelasyonu iki-yarı güvenilirliğinin bir ölçüsü olarak hesaplarız. Enstrüman ne kadar uzun olursa, ölçünün iki yarısının benzer olması o kadar olabirlikli olacaktır (daha çok madde ilave edildikçe, rastgele hatalar minimize olacağı için) ve böylece bu teknik daha uzun enstrümanları sistematik olarak yüksek-kestirmeye eğilimli olacaktır.

İçsel tutarlılık güvenilirliği. İçsel tutarlılık güvenilirliği aynı yapının farklı maddeleri arasındaki tutarlılığın bir ölçüsüdür. Eğer çok-maddeli bir yapı ölçüsü yanıtlayıcılara uygulanırsa, yanıtlayıcıların o maddeleri benzer bir tarzda reytinglendirmelerinin düzeyi içsel tutarlılığın bir yansımasıdır. Bu güvenilirlik ortalama maddeler-arası korelasyon, ortalama madde-total korelasyonu veya daha müştereken Cronbach'ın alfası cinsinden kestirilebilir. Bir misal vermek icabederse, eğer altı maddeli bir ölçeğimiz varsa, bu altı madde arasında on beş farklı madde çiftine ve on beş korelasyona sahip oluruz. Ortalama maddeler-arası korelasyon bu on beş korelasyonun ortalamasıdır. Ortalama madde-total korelasyonunu hesaplamak için, altı maddenin tümüne ait değerleri toplayarak bir "total" madde yaratmamız ve bu total madde ve münferit maddelerin her biri arasındaki korelasyonları hesaplamamız ve en sonunda bu altı korelasyonun ortalamasını almamız gerekir. Yukarıdaki iki ölçünün hiç biri ölçüdeki maddelerin sayısını (bu misalde altı madde) hesaba katmaz. Cronbach'ın alfası, Lee Cronbach tarafından 1951 de dizayn edilmiş olan bir güvenilirlik ölçüsü olarak, güvenilirlik kestiriminde ölçek genişliğini bir faktör olarak hesaba katar, şöyle ki:

$$\alpha = \frac{K}{K-1} \left(1 - \frac{\sum_{i=1}^K \sigma_{Y_i}^2}{\sigma_X^2} \right)$$

ki burada K ölçüdeki maddelerin sayısı, σ_X^2 gözlenen total skorların varyansı (standart sapmanın karesi) ve $\sigma_{Y_i}^2$ i . madde için gözlenen varyanstır. Standartlaştırılmış Cronbach'ın alfası daha basit bir formül kullanılarak hesaplanabilir:

$$\alpha_{\text{standartlaştırılmış}} = \frac{K\bar{r}}{(1 + (K-1)\bar{r})}$$

ki burada K maddelerin sayısı, $\bar{r}$ ortalama maddeler-arası korelasyon, yani, korelasyon matrisinin üst (veya alt) üçgenselindeki $K(K-1)/2$ katsayının aritmetik ortalamasıdır.

Geçerlilik

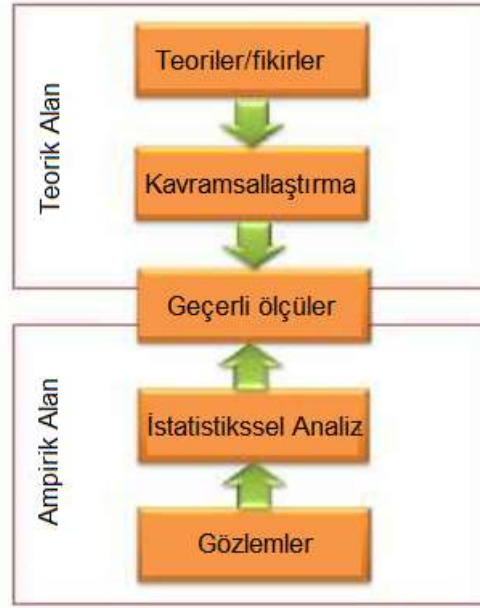
Geçerlilik (/validity), çoğu zaman yapı geçerliliği olarak adlandırılır, bir ölçünün ölçtüğü

farzedilen alttayatan yapıyı ne kadar kifayyetle temsil ettiğini ifade eder. Mesela, bir şefkat ölçüsü, mesela empati gibi başka bir yapıyı değil de, gerçekten şefkati mi ölçüyor? Geçerlilik teorik ve ampirik yaklaşımları kullanmak suretiyle değerlendirilebilir. Geçerliliğin teorik değerlendirmesi bir teorik yapı fikrinin bir operasyonel ölçüye ne kadar iyi çevrilebildiğine veya bir operasyonel ölçü ile ne kadar iyi temsil edilebildiğine odaklanır. Bu tip geçerlilik **çevirisel geçerlilik** (veya temsilsel geçerlilik) olarak adlandırılır ve iki alt-tipten meydana gelir: yüz geçerliliği ve içerik geçerliliği. Çevirisel geçerlilik her bir maddeyi (indikatör) o yapının kavramsal tanımını ile ne kadar iyi uyumlu olduğuna göre reytinglendiren uzman yargıcıların bir paneli ve Q-sort adı verilen nitel bir teknik kullanılarak değerlendirilir, tipik olarak.

Geçerliliğin ampirik değerlendirmesi belirlenen ölçülerin bir veya daha çok sayıdaki dışsal kriterle ne kadar iyi ilişkili olduğunu ampirik gözlemlere dayanarak inceler. Bu tip geçerlilik **kriter-ilişkili geçerlilik** olarak adlandırılır, ki dört alt-tipi vardır: yakınsak, ayırmsal, eş-akışlı ve tahminsel geçerlilik. Çevirisel geçerlilik bir ölçünün altında yatan yapıyı ne kadar iyi yansıttığını incelerken, kriter-ilişkili geçerlilik belirli bir ölçünün, o yapının teorisine göre davranması gerektiği üzere davranıp davranmadığını inceler. Bu değerlendirme gözlenen verinin mesela korelasyon analizi, faktör analizi ve benzeri istatistiksel teknikler kullanılarak analizine dayanır. Geçerliliğin teorik ve ampirik değerlendirmesi arasındaki ayrım Figür 7.2 de resmedilmektedir. Mamafih, her iki yaklaşımın da sosyal bilim araştırmasında ölçülerin geçerliliğini kifayyetle sağlaması gerekir.

Dikkat etmelisin ki geçerliliğin burada tartışılan farklı tipleri *ölçüm prosedürlerinin* geçerliliğine matuftur, ki bu mesela içsel geçerlilik (nedensellik), dışsal geçerlilik (genelleştirilebilirlik) gibi *hipotez testi prosedürlerinin* geçerliliğinden veya istatistiksel netice geçerliliğinden ayrıdır. Geçerliliğin sonraki tipleri daha sonraki bir bölümde tartışılmaktadır.

Yüz geçerliliği. Yüz geçerliliği bir indikatörün altında yatan yapının “yüzeyde” makul bir ölçüsü olarak görünüp görünmediğine matuftur. Mesela, birisinin müşterek dini ritüellere katılımının frekansı o kişinin dindarlığının bir indikatörü olarak, çok açıklama olmaksızın, manalı görünür. Mamafih, eğer çalışan moralinin bir ölçüsü olarak, bir ofis kütüphanesinden kaç kitabın ödünç alındığını teklif edersek, o zaman böyle bir ölçünün muhtemelen yüz geçerliliği eksik olacaktır çünkü pek manalı görünmemektedir. İlginçtir ki, organizasyonel araştırmada göze çarpan bazı popüler ölçülerin yüz geçerliliği eksiktir. Mesela, bir organizasyonun absortif kapasitesi (organizasyonel prosesleri için ne kadar yeni bilgi asimile edebildiği) çoğu zaman araştırma ve geliştirme yoğunluğu (yani, R&D harcamaları bölü toplam gelir) olarak ölçülmektedir! Eğer araştırmamız hayli soyut yapıları veya kavramsal olarak birbirinden ayrılması zor olan yapıları (mesela, şefkat ve empati gibi) ihtiva ediyorsa, yapı ölçülerimizin yüz geçerliliğini değerlendirmek üzere bir uzmanlar paneli kullanmayı düşünmek yararlı olabilir.



Figür 7.2 Geçerlilik değerlendirmesinin iki yaklaşımı

İçerik geçerliliği. İçerik geçerliliği bir ölçek maddeleri kümesinin ölçülmeye çalışılan yapının münasebetli içerik alanı ile ne kadar iyi eşleştiğinin bir değerlendirmesidir. Mesela, “restoran hizmetinden memnuniyet” yapısını ölçmek istersen ve restoran hizmetinin içerik alanını yiyecek kalitesi, hizmet elemanlarının nezaketi, bekleme süresi ve restoranın genel ambiyansı (yani, gürültülü, dumanlı vb. olup olmaması) olarak tanımlarsan, o zaman kifayyetli içerik geçerliliği için, yiyecek kalitesinden, hizmet elemanlarının nezaketinden, bekleme süresinden ve restoranın ambiyansından bir müşterinin ne kadar memnun kaldığını inceleyen indikatörler kullanılarak bu yapı ölçülmelidir. Elbette, bu yaklaşım bir yapının bütün içerik alanının detaylı bir tanımını gerektirir, öyle ki mesela özgüven veya zeka gibi kompleks yapılar için bu zor olabilir. Bu yüzden, içerik geçerliliğini kifayyetle değerlendirmek her zaman olanaklı olmayabilir. Yüz geçerliliğinde olduğu üzere, yargıcıların bir uzman paneli, yapıların içerik geçerliliğini değerlendirmek için istihdam edilebilir.

Yakınsak geçerlilik bir ölçünün ölçmeyi kastettiği yapı ile ilişkisinin yakınlığına (veya yakınsamasına) atıfta bulunur ve **ayrımsal geçerlilik** bir ölçünün ölçmeyi kastetmediği başka yapıları ne derece ölçmediğine (veya ayrımsamasına) atıfta bulunur. Çoğunlukla, yakınsak geçerlilik ve ayrımsal geçerlilik, ilişkili yapıların bir kümesi için, birlikte değerlendirilir. Mesela, eğer bir organizasyonun bilgisinin performansı ile ilişkili olması beklenen bir durum ise, organizasyonel bilgi ölçümümüzün gerçekten organizasyonel bilgiyi ölçtüğünden (yakınsak geçerliliğinden) ve organizasyonel performansı ölçmediğinden (ayrımsal geçerliliğinden) nasıl emin olabiliriz? Yakınsak geçerlilik, bir yapının bir indikatörünün gözlenen değerlerini aynı yapının diğer indikatörlerinininkiler ile karşılaştırmak ve bu indikatörler arasındaki benzerliği (veya yüksek korelasyonu) göstermek suretiyle tespit edilebilir. Ayrımsal geçerlilik bir yapının indikatörlerinin diğer yapılar ile benzer olmadığını (yani düşük korelasyonlu olduğunu) göstermek suretiyle tespit edilebilir. Yukarıdaki misalde, eğer organizasyonel bilginin üç-maddeli bir ölçüsüne ve organizasyonel performansın yine üç-maddeli bir ölçüsüne sahipsek, bilgi ve performans maddelerinin her bir çifti arasında ikideğişkeli korelasyonları hesaplayabiliriz. Eğer bu korelasyon matrisi organizasyonel bilgi ve organizasyonel performans yapılarının kendi maddeleri içerisinde yüksek korelasyonlar gösteriyor fakat bu yapılara ait maddeler arasında düşük korelasyonlar gösteriyorsa, o zaman yakınsak geçerliliği ve ayrımsal geçerliliği eşzamanlı olarak göstermiş oluruz (Tablo 7.1 e bak).

	KL1	KL2	KL3	PF1	PF2	PF3
KL1	1.00	0.83	0.79	0.23	0.21	0.19
KL2		1.00	0.75	0.11	0.20	0.03
KL3			1.00	0.03	-0.11	0.17
PF1				1.00	0.84	0.91
PF2					1.00	0.77
PF3						1.00

Aynı yapının maddeleri arasında yüksek korelasyonlar (yakınsak geçerlilik)

Farklı yapıların maddeleri arasında düşük korelasyonlar (ayrimsal geçerlilik)

Tablo 7.1 Yakınsak ve ayrimsal geçerlilik için iki-değişkeli korelasyon analizi

Yakınsak ve ayrimsal geçerlilik için alternatif ve daha müşterek bir istatistiksel metot *keşifsel faktör analizidir*. Bu bir veri indirgeme tekniğidir ki ele alınan maddeler kümesini, başlıca bileşenler analizi adı verilen bir istatistik tekniğini kullanmak suretiyle, yukarıda tartışılan ikideğişkeli korelasyon yapısına dayanarak daha küçük bir faktörler kümesinde toplastırır. Bu faktörler ölçmeye çabaladığımız altta-yatan teorik yapılara tekabül edecektir, ideal olarak. Faktör ekstraksiyonu için genel norm ekstrete edilen her bir faktörün 1.0 dan daha büyük bir özdeğere sahip olmasıdır. Ekstrete edilen faktörler, daha sonra, her bir yapının bireysel maddelerini bir kompozit ölçü içerisinde toplastırmakta kullanılabilecek faktör ağırlıklarının hesaplamak için, altta-yatan yapıların nispeten korelasyonsuz veya korelasyonlu olmalarının beklenip beklenmediğine bağlı olarak, ortogonal veya oblik rotasyon teknikleri kullanılarak rotasyona tabi tutulabilirler. Kifayetli yakınsak geçerlilik için, müşterek bir yapıya ait maddelerin tek bir faktörde 0.60 veya daha yüksek faktör yüklerine (aynı-faktör yükleri olarak adlandırılır) sahip olmaları beklenirken, ayrimsal geçerlilik için, bu maddelerin diğer tüm faktörlerde 0.30 veya daha az faktör yüklerine (çapraz-faktör yükleri olarak adlandırılır) sahip olmaları gerekir; Tablo 7.2 deki rotasyona uğratılmış faktör matrisi misalinde görüldüğü üzere. Yakınsak geçerliliği ve ayrimsal geçerliliği değerlendirmek için daha sofistike bir teknik çoklu-vasıf çoklu-metot (MTMM) yaklaşımıdır. Bu teknik her yapıyı (vasıf) iki veya daha çok sayıda metot kullanarak ölçmeyi gerektirir (mesela, yoklama veya kişisel gözlem; veya akademik kaliteyi değerlendirmek için mesela öğretmenler ve ebeveynler gibi iki farklı yanıtlayıcı grubunu belki yoklamak). Bu meşakkatli ve nispeten daha az popüler olan bir yaklaşımdır ve bu yüzden burada detaylıca tartışılmayacaktır.

	Faktör 1	Faktör 2
KL1	0.88	0.13
KL2	0.93	0.11
KL3	0.87	0.03
PF1	0.17	0.93
PF2	-0.03	0.85
PF3	0.07	0.78

Yüksek aynı-faktör yükleri (yakınsak geçerlilik)

Düşük çapraz-faktör yükleri (ayrimsal geçerlilik)

Tablo 7.2. Yakınsak ve ayrimsal geçerlilik için faktör analizi.

Kriter-ilişkili geçerlilik aynı zamanda belirli bir ölçünün güncel veya müstakbel bir kriter ile iyi ilişkili olup olmadığına dayanarak ta değerlendirilebilir, ki bu değerlendirme sırasıyla eş-akışlı geçerlilik ve tahminsel geçerlilik olarak adlandırılır. **Tahminsel geçerlilik** bir ölçünün teorik olarak beklenen müstakbel bir çıkımı başarıyla tahmin etme derecesidir. Mesela, standartlaştırılmış test skorları (yani, üniversiteye giriş sınavı skorları) üniversitede akademik başarıyı (yani, mezuniyet not ortalaması) doğru olarak tahmin edebilir mi? Böyle bir geçerliliği değerlendirmek yapıların birbirleriyle teorik olarak nasıl ilişkili olduklarını gösteren bir “nomolojik ağı” yaratılmasını gerektirir. **Eş-akışlı geçerlilik** bir ölçünün eş-zamanlı neydan geldiği farzedilen başka bir somut kriterle ne kadar iyi ilişkili olduğunu inceler. Mesela, öğrencilerin analiz dersinden aldıkları sınav skorları cebir dersinden aldıkları sınav skorları ile ne kadar korelasyonludur? Bu skorların eş-akışlı olarak ilişkili olmaları gerekir çünkü her ikisi de matematik testidir. Yakınsak ve ayrımsal geçerliliğin tersine, eş-akışlı ve tahminsel geçerlilik ampirik sosyal bilim araştırmasında çoğu zaman gözardı edilir.

Ölçüm Teorisi

Madem ki güvenilirliğin ve geçerliliğin farklı çeşitlerini biliyoruz, şimdi, güvenilirlik ve geçerlilik anlayışımızı, aynı zamanda *hakiki skor teorisi* olarak ta adlandırılan *klasik test teorisini* kullanmak suretiyle, matematiksel yoldan sentezlemeye çalışalım. Bu teori, ölçümün nasıl işlediğini, neyi ölçtüğünü ve neyi ölçmediğini inceleyen bir psikometrik teoridir. Bu teori şunu postüla eder ki her gözlem bir *hakiki skora* (T) sahiptir ki bu puan isabetlice gözlenebilir eğer ölçümde hata yoksa. Mamafi, *ölçüm hatasının* (E) mevcudiyeti *gözlenen skor* (Y) ve hakiki skor arasında bir sapma ile sonuçlanır:

$$\begin{array}{ccccc} Y & = & T & + & E \\ \text{Gözlenen skor} & & \text{Hakiki skor} & & \text{Hata} \end{array}$$

Gözlenen skorların bir kümesinde, gözlenen ve hakiki skorların varyansları benzer bir eşitlik vasıtasıyla ilişkilendirilebilir:

$$\text{var}(Y) = \text{var}(T) + \text{var}(E)$$

Psikometrik analizin amacı hata varyansını ($\text{var}(E)$), eğer olanaklıysa, kestirmek (*estimate*) ve minimize etmektir, öyle ki gözlenen Y skoru hakiki T skorunun iyi bir ölçüsü olsun.

Ölçüm hataları iki tipte olabilir: rastgele hata ve sistematik hata. **Rastgele hata** öyle bir hatadır ki gözlemlerin bazılarında rastgele tesir eden, bilinmeyen ve kontrol edilebilmeyen dışsal faktörlerin bir kümesine hamledilebilir. Mesela, ölçüm yapılırken, bazı yanıtlayıcılar diğerlerine göre daha nezaketli bir modda olabilirler ki bu da ölçüm maddelerine verecekleri yanıtlara tesir edebilir. Mesela, daha nezaketli bir moddaki yanıtlayıcılar özgüven, memnuniyet ve mutluluk gibi yapılara, kaba bir modda olanlara göre daha pozitifçe yanıt verebilirler. Mamafi, araştırma öğrenme-çalışmalarında, hangi deneğin hangi modda olduğunu önceden anlamak veya modun etkisini kontrol altına almak olanaklı değildir. Aynı minval üzere, organizasyonel bir düzeyde, eğer firma performansını ölçüyorsak, idari veya çevresel değişiklikler gözlenen bir örnekteki firmaların bazılarının performanslarını etkileyebilir fakat diğerlerinininkini değil. Bundan ötürü, ölçmede, rastgele hata “parazit” olarak addedilir ve genellikle gözardı edilir.

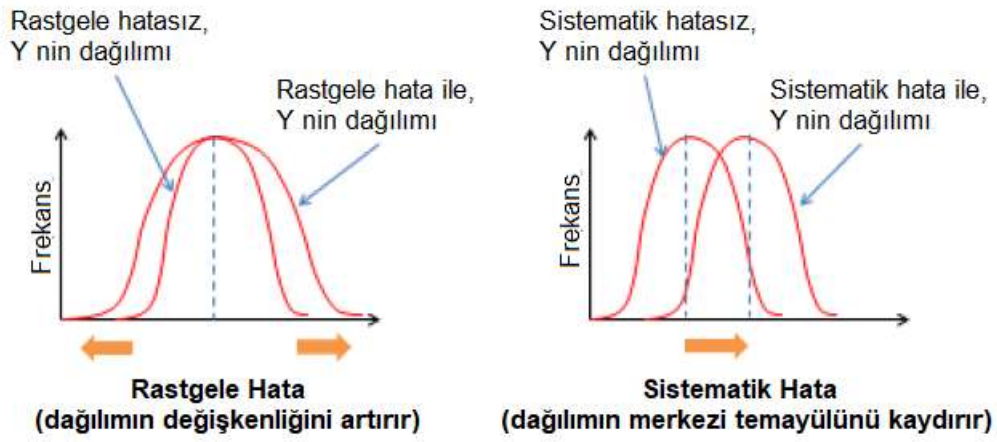
Sistematik hata öyle bir hatadır ki bir yapının bir örnekteki tüm gözlemlerini sistematik olarak etkileyen faktörler tarafından meydana getirilir. Bir önceki firma performansı misalimizde, farzedelim ki yakın zamanda çıkan bir finansal kriz mesela finans firmalarının performansına, imalat veya hizmet firmaları gibi firma tiplerine göre daha fazla tesir etmiş olsun. Eğer bizim örneğimiz yalnızca finans firmalarından oluşuyor ise, örneğimizdeki tüm firmaların

performanslarında, finansal krizden dolayı, sistematik bir düşüklük bekleyebiliriz. Rastgele hata, bir örnekteki gözlemlerde pozitif, negatif veya sıfır olabildiği halde, sistematik hatalar, örneğin tümünde mütemediyen pozitif veya negatif olmak eğilimindedir. Bundan ötürü, sistematik hata, ölçmede, bazen, “yanlılık” olarak mülahaza edilir ve düzeltilmesi gerekir.

Gözlenen bir skor rastgele ve sistematik hataların her ikisini de ihtiva edebileceği için, hakiki skor eşitliğimiz şöyle modifiye edilebilir:

$$Y = T + E_r + E_s$$

ki burada E_r ve E_s sırasıyla rastgele ve sistematik hataları temsil eder. Bu hataların istatistiksel tesiri şudur ki rastgele hata gözlenen bir ölçünün dağılımına değişkenlik (yani, standart sapma) kazandırır fakat onun merkezi temayülüne (yani, aritmetik-ortalamasına) etki etmezken, sistematik hata merkezi temayülü etkiler fakat değişkenliği değil (bak, Figür 7.3).



Figür 7.3. Rastgele ve sistematik hataların etkileri.

Ölçüm prosesi için, rastgele ve sistematik hata ne manaya gelir? Gözlemlerdeki değişkenliği artırmak suretiyle, rastgele hata ölçümün güvenilirliğini azaltır. Aksine, merkezi temayülü kaydırmak suretiyle, sistematik hata ölçümün geçerliliğini azaltır. Geçerlilik kaygıları, güvenilirlik kaygılarına göre, ölçmede çok daha ciddi problemlerdir. Çünkü geçersiz bir gözlem ölçülmek istenilenden farklı bir yapıyı ölçüyor olabilir ve bundan ötürü, geçerlilik problemleri istatistiksel analizden türetilen bulgulara ciddi şüphelere neden olur.

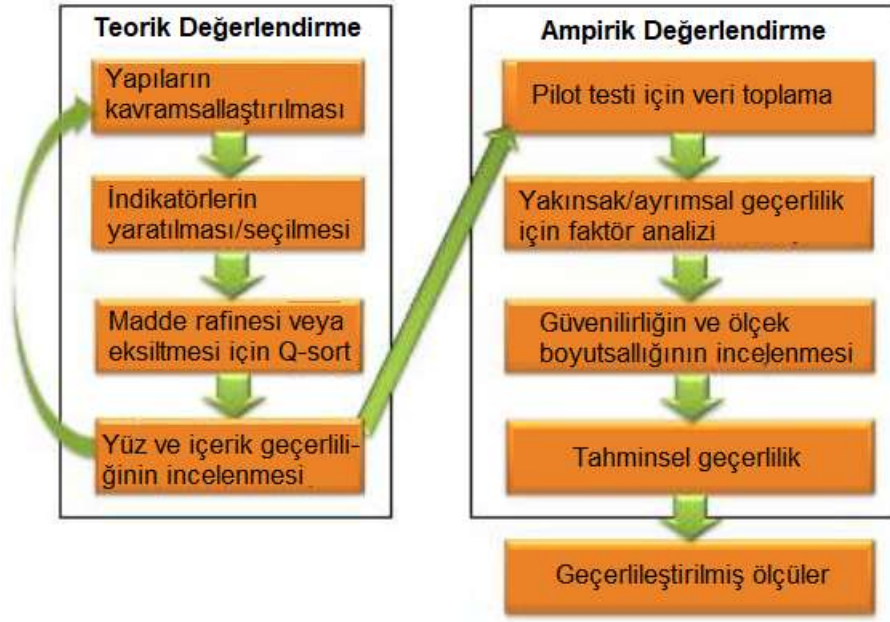
Dikkat etmelisin ki güvenilirlik bir oran veya kesirdir ki hakiki skorun gözlenen skora nispi yakınlığını ifade eder. Bu manada, güvenilirlik şöyle ifade edilebilir:

$$\frac{\text{var}(T)}{\text{var}(Y)} = \frac{\text{var}(T)}{\text{var}(T) + \text{var}(E)}$$

Eğer $\text{var}(T) = \text{var}(Y)$ ise, o zaman hakiki skor gözlenen skorla aynı değişkenliğe sahiptir ve güvenilirlik 1 dir.

Ölçüm Geçerliliğine Entegre Bir Yaklaşım

Geçerliliğin tam ve kifayetli bir değerlendirmesi hem teorik hem de ampirik yaklaşımları ihtiva etmelidir. Figür 7.4 te gösterildiği üzere, bu detaylı ve çok-adımlı bir prosestir ki farklı ölçek güvenilirliği ve geçerliliği tiplerinin hesaba katılmasını gerektirir.



Figür 7.4. Ölçüm geçerliliğine entegre bir yaklaşım.

Entegre yaklaşım teorik alanda başlar. Birinci adım ilgilenilen yapıların kavramsallaştırılmasıdır. Bu her yapıyı tanımlamayı ve onları oluşturan alanları ve/veya boyutları teşhis etmeyi ihtiva eder. Sonrasında, her yapı için, bu yapıların kavramsallaştırılmasına dayanan maddeler veya indikatörler, Bölüm 6 daki ölçeklendirme prosedüründe betimlendiği üzere seçilir. İndikatör seçiminde bir literatür taraması yardımcı olabilir. Her madde, basit ve anlaşılması kolay bir dil kullanılarak, üniform bir usupla kelimelere dökülür. Bu adımı takiben, uzman yargıcıların bir paneli (araştırma metodlarında uzman akademisyenler ve/veya hedef yanıtlayıcıların temsilci bir kümesi), her bir indikatörü incelemek ve bir Q-sort analizi icra etmek için, istihdam edilebilir. Bu analizde, kavramsal tanımlarıyla beraber tüm yapıların bir listesi ve yapı ölçülerinin her birine ait indikatörleri listeleyen endeks kartları (bir endeks kartına bir indikatör) her yargıcıya verilir. Yargıcılardan, sonrasında, her bir endeks kartını bağımsızca okumaları, o maddenin açıklığını, okunabilirliğini ve semantik manasını incelemelerini ve kendilerine verilen yapı tanımlarına dayanarak, o maddeyi içerisinde en fazla manalı buldukları yapıya dahil etmeleri istenir. Yargıcıların kendi sınıflandırmaları ile ne kadar mutabık olduklarını incelemek için, reytingciler-arası güvenilirlik değerlendirilir. Yargıcıların çoğu tarafından müştereken ıskalanan muğlak maddeler yeniden incelenebilir, yeniden ifade edilebilir veya atılabilir. İleri analiz için, her yapıya ait en iyi maddeler (diyelim 10-15) seçilir. Seçilen maddelerin her biri yargıcılar tarafından, yüz geçerliliği ve içerik geçerliliği bakımından yeniden incelenir. Eğer bu safhada maddelerin kifayetli bir kümesi temin edilemez ise, kastedilen yapının kavramsal tanımına dayanan yeni maddeler yaratmak gerekebilir. İlgilenilen yapıları en iyi temsil eden maddelerin bir kümesi üzerinde yargıcılar arasında makul bir anlaşma sağlamak için Q-sort analizinin iki veya üç turuna ihtiyaç duyulabilir.

Sonrasında, geçerleme prosedürü ampirik alana taşınır. Rafine edilen yapı maddelerinin tümü bir araya getirilerek, bir araştırma enstrümanı yaratılır ve hedef popülasyonun temsilci yanıtlayıcılarının bir pilot test grubuna uygulanır. Toplanan veri organize edilir ve yakınsak ve ayrımsal geçerliliği değerlendirmek için, mesela SAS veya SPSS gibi yazılım programları kullanılarak, korelasyonel analize veya keşifçi faktör analizine tabi tutulur. Beklenen faktör yükü normlarını (aynı-faktör yüklerinin 0.60 tan büyük olması ve çapraz-faktör yüklerinin 0.30 dan küçük olması) karşılamayan maddeler bu safhada atılmalıdır. Elde kalan ölçekler,

güvenilirlik için, mesela Cronbach'ın alfası gibi bir içsel tutarlılık ölçüsü kullanılarak, değerlendirilir. Bu safhada, hedeflenen yapıların tek-boyutlu olarak mı yoksa çok-boyutlu olarak mı kavramsallaştırıldıklarına bağlı olarak, ölçek boyutsallığı da doğrulanabilir. Sonrasında, her bir yapının tahminsel kabiliyeti, regresyon analizi veya sütrüktürel eşitlik modellemesi kullanılarak, yapının teorik olarak spesifikleştirilen bir nomolojik ağı içerisinde değerlendirilir. Eğer yapı ölçüleri bu bölümde betimlenen güvenilirlik ve geçerlilik gerektirimlerinin tümünü veya çoğunu karşılıyor ise, opersyonelleştirilen ölçülerin makul surette kifayetli ve isabetli olduklarından emin olunabilir.

Burada tartışılan ölçüm geçerliliğine ait entegre yaklaşım araştırmacıdan hayli zaman ve çaba ister. Buna rağmen, araştırmamızda kullanacağımız ölçeklerin bilimsel araştırmanın beklenen normlarını karşıladığından emin olmak için, bu detaylı ve çok-adımlı prosese gereksinim duyarız. Kusurlu veya zayıf ölçekler kullanılarak varılan neticeler manasız olduğu için, ölçek geçerkleme ve ölçme ampirik araştırmanın en önemli ve girift fazlarından birini teşkil eder.

Bölüm 8

Örnekleme

Örnekleme ilgilenilen bir popülasyonun bir alt-kümesinin (ki “örnek” olarak adlandırılır) seçimine matuf istatistiksel procestir. Örneklemenin maksadı gözlemler yapmak suretiyle ilgilenilen popülasyon hakkında istatistiksel çıkarımlarda bulunmaktır. Sosyal bilim araştırması genellikle spesifik popülasyonlara ait davranış biçimleri hakkında çıkarımlar yapmaya daırdır. Fizibilite ve maliyet kısıtları popülasyonun tamamını öğrenmeye-çalışmaya çoğu zaman engel olur ve bundan ötürü, gözlem ve analiz için, ilgilendiğimiz popülasyondan bir temsilci örnek seçmemiz gerekir. Popülasyonun hakikaten temsilcisi olan bir örnek seçmek aşırı önem arzeder öyle ki örnekten türetilen çıkarımlar ilgilenilen popülasyona geri genelleştirilebilsin. Münasip olmayan ve yanlış örnekleme mesela CNN/Gallup Poll, ABC ve CBS gibi farklı yoklama şirketleri tarafından her Amerika birleşik Devletleri Başkanlık seçiminden önce yapılan halk oyu yoklamaları ve sandık-çıkışı yoklamalarında rapor edilen ıraksak ve hatalı çıkarımların birincil sebebidir.

Örnekleme Prosesi

Örnekleme prosesi çeşitli safhalardan oluşur. Birinci safha hedef popülasyonu tanımlamaktır. Bir **popülasyon** bir araştırmacının öğrenmeye-çalışmak istediği karakteristiklere sahip insanların veya bireylerin (**analiz ünitesi**) tümü olarak tanımlanabilir. Analiz ünitesi bir kişi, grup, organizasyon, ülke, nesne veya hakkında bilimsel çıkarımlarda bulunmayı arzu edeceğimiz başka bir entite olabilir. Bazen popülasyon aşıkardır. Mesela, eğer bir imalatçı bir üretim bandında imal edilen mamüllerin belli kalite gereksinimlerini karşılayıp karşılamadığını belirlemek istiyorsa, o zaman popülasyon o üretim bandında imal edilen mamüllerin tümünden meydana gelir. Diğer zamanlarda, hedef popülasyonu anlamak biraz daha zor olabilir. Eğer lise öğrencileri arasında akademik öğrenmenin başlıca saiklerini teşhis etmek istiyorsak, o zaman hedef popülasyonumuz ne olurdu: lise öğrencileri mi, öğrencilerin öğretmenleri mi, okul müdürleri mi yoksa öğrencilerin ebeveynleri mi? Bu durumda doğru cevap lise öğrencileridir. Çünkü onların performansı ile ilgileniyoruz, onların öğretmenlerinin, ebeveynlerinin veya okulların değil. Aynı minvalde, eğer yanlış çarkları teşhis etmek için, rulet çarklarının davranışını analiz etmek istiyorsak, ilgilendiğimiz popülasyon tek bir rulet çarkının farklı gözlemleri değil, fakat farklı rulet çarklarıdır (yani, çarkların sonsuz bir kümesindeki davranış).



Figür 8.1. Örnekleme prosesi.

Örnekleme prosenin ikinci safhası bir **örnekleme çerçevesi** seçmektir. Bu, örneğin fiilen içinden çekilebileceği hedef popülasyonun ulaşılabilir bir kısmıdır (genellikle iletişim enformasyonu içeren bir liste). Eğer hedef popülasyonumuz istihdam edilen profesyoneller ise, dünya üzerinde istihdam edilen tüm profesyonellere ulaşabilmeyeceğimiz için, daha gerçekçi bir örnekleme çerçevesi bizim öğrenme-çalışmamızda yer almaya istekli bir veya iki yerel şirketin eleman listeleri olacaktır. Eğer hedef popülasyonumuz organizasyonlar ise, o zaman firmaların Fortune 500 listesi veya New York Borsasında kayıtlı firmaların Standart & Poor listesi kabul edilebilir örnekleme çerçeveleri olabilirler.

Dikkat etmelisin ki örnekleme çerçeveleri popülasyonun hepsini temsil ediyor olmayabilirler ve eğer böyle ise, böylesi bir örnekten türetilen çıkarımlar popülasyona genelleştirilebilir olmayabilir. Mesela, eğer hedef popülasyonumuz organizasyonel çalışanların hepsi ise (mesela, bu popülasyonda çalışanların özgüvenini öğrenmeye-çalışmayı arzu ediyor olabiliriz) ve örnekleme çerçevemiz Amerikan ortabatısındaki otomotiv şirketlerinin çalışanları ise, bu gruplardan elde edilecek bulgular, bırakın küresel işyerlerini, Amerikan işgücüne dahi genelleştirilebilir olmayabilir. Bunun sebebi şu ki Amerikan otomotiv endüstrisi son 50 yıldır şiddetli rekabet baskısı altındadır ve yaşamış olduğu çok sayıdaki yeniden organize edilme ve küçülme epizodlarının daha düşük çalışan morali ve özgüveni ile sonuçlanmış olması mümkündür. Dahası, Amerikan işgücünün çoğunluğu hizmet endüstrilerinde veya küçük işletmelerde istihdam edilmektedir, otomotiv endüstrisinde değil. Bundan ötürü, Amerikan otomotiv endüstrisi çalışanlarının bir örneği Amerikan işgücünün temsilcisi değildir. Aynı minvalde, Fortune 500 listesi en büyük 500 Amerikan firmasını ihtiva eder ki pek çoğu büyükten ziyade orta ve küçük hacimli olan tüm Amerikan firmalarının genelde temsilcisi değildir ve bu yüzden yanlış bir örnekleme çerçevesidir. Aksine, S&P listesi, S&P large-cap, mid-cap veya small-cap listelerini kullanmak suretiyle, büyük, orta ve/veya küçük şirketleri seçmemize olanak sağlar fakat yalnızca halka açık şirketleri ihtiva ettiği için, yine yanlıdır. Aynı zamanda dikkatini çekmek istiyorum ki bir örneğin içinden çekildiği popülasyon illa ki hakkında enformasyon elde etmek istediğimiz popülasyonla birebir aynı olmak zorunda değildir. Mesela, eğer bir araştırmacı yeni bir “sigara bıraktırma” programının başarı oranını öğrenmek istiyorsa, o zaman hedef popülasyon bu programa katılma olanağına sahip içicilerin evrenidir ki belki de bu sonsuz veya meçhul bir popülasyondur. Bundan ötürü, araştırmacı

sigara bırakma tedavisi için yerel bir sağlık kurumuna gelen hastaları örnekleyebilir ki bu hastaların bazıları bu belirli “sigara bıraktırma” programına maruz bırakılmayabilir. Bu durumda örnekleme çerçevesi ilgilenilen popülasyon ile tam olarak örtüşmez.

Örneklemenin son safhası, iyi-tanımlı bir örnekleme tekniği kullanmak suretiyle, örnekleme çerçevesinden bir örnek seçmektir (veya çekmektir). Örnekleme teknikleri iki geniş kategori içerisinde sınıflandırılabilir: olasılıklı (rastgele) örnekleme ve olasılıksız örnekleme. Olasılıklı örnekleme idealdir eğer sonuçların genelleştirilebilirliği öğrenme-çalışmamız için önemliyse. Fakat kendine mahsus bir takım koşullar altında olasılıksız örneklemenin kullanılması da gerekçelendirilebilir. Bu teknikleri müteakip iki alt-bölümde tartışacağız.

Olasılıklı Örnekleme

Olasılıklı örnekleme öyle bir teniktir ki bu teknikle popülasyondaki her ünitenin örneğe seçilmek için bir şansı (sıfırdan-farklı bir olasılık) vardır ve bu şans isabetlice belirlenebilir. Bu sayede, üretilen mesela örnek aritmetik-ortalaması veya standart sapma gibi örnek istatistikleri, örnek üniteleri seçilme olasılıklarına göre ağırlıklandırıldıkları sürece, popülasyon parametrelerinin yansız kestirimleridir. Tüm olasılıklı örneklemeler müşterek iki vasfı havidir: (1) popülasyondaki her ünite bilinen sıfırdan-farklı bir örneklenme olasılığına sahiptir ve (2) örnekleme prosedürü rastgele seçimi bir noktada ihtiva eder. Olasılıklı örneklemenin muhtelif tipleri şunlardır:

Basit rastgele örnekleme. Bu teknikte, popülasyonun (daha isabetlice, bir örnekleme çerçevesinin) tüm olanaklı alt-kümeleri eşit bir seçilme şansına sahiptir. Bir örnekleme çerçevesindeki total N ünite içerisinde n ünitelik herhangi bir kümeyi seçme olasılığı C_n^N dir. Bundan ötürü, örnek istatistikleri, herhangi bir ağırlıklandırma yapılmaksızın, popülasyon parametrelerinin yansız kestirimleridir. Basit rastgele örnekleme yanıtlayıcıları örnekleme çerçevesinden rastgele seçmeyi gerektirir. Fakat geniş örnekleme çerçeveleri söz konusu olduğu zaman, bir rastgele sayılar tablosu veya kompütörde bir rastgele sayı üreticisi kullanılır. Mesela, eğer 1000 firmayı barındıran bir listeden, yoklamak için, 200 firmayı seçmek istiyorsak, ve eğer bu liste mesela Excel gibi bir veri-sayfasına girilmişse, listedeki 1000 firmanın her biri için rastgele sayılar üretmek üzere, Excel’in RAND() fonksiyonunu kullanabiliriz. Daha sonra, listeyi müteakip rastgele sayılara göre küçükten büyüğe doğru sıralarız ve bu sıralı listeden ilk 200 firmayı seçeriz. Tüm olasılıklı örnekleme tekniklerinin, bu en basitidir; ancak, basitliği bu tekniğin, aynı zamanda gücüdür de. Örnekleme çerçevesi alt-bölgelere ayrılmadığı ve parçalanmadığı için, örnek yansızdır ve tüm olasılıklı örnekleme teknikleri arasında çıkarımlar en genelleştirilebilirdir.

Sistematik örnekleme. Bu teknikte, örnekleme çerçevesi bazı kriterlere göre sıralanır ve sıralı liste boyunca, düzenli aralıklarla, üniteler seçilir. Sistematik örnekleme rastgele bir başlangıç noktası tespit etmeyi gerektirir ve sonrasında bu noktadan itibaren her k inci ünitenin seçimi ile devam eder ki burada $k = N/n$, yani örnekleme çerçevesinin hacmi, N in arzulanan örnek hacmi n e oranıdır ve k formel olarak *örnekleme oranı* olarak adlandırılır. Başlangıç noktasının otomatik olarak listenin birinci elemanı olmaması önemlidir. Bunun yerine, listenin ilk k elemanı arasından rastgele seçilmelidir. Bin firmadan oluşan bir listeden 200 firmanın seçildiği hemen önceki misalimizde, 1000 firmayı büyüklüklerine göre (yani, çalışan sayısı veya yıllık gelir) küçükten büyüğe (veya büyükten küçüğe) doğru sıralayabiliriz, sıralı listedeki ilk 5 firmadan birini rastgele seçebiliriz ve daha sonra listedeki her beşinci firmayı seçebiliriz. Bu proses şundan emin olmamızı temin edecektir ki örneğimizde daha büyük veya daha küçük firmalar aşkın temsil edilmeyecektir fakat bundan ziyade tüm büyüklüklerdeki firmalar, genel olarak örnekleme çerçevemizde olduğu gibi temsil edileceklerdir. Başka bir deyişle, örnek popülasyonun temsilcisi olacaktır, en azından sıralama kriterine istinaden.

Tabakalandırılmış örnekleme. Tabakalandırılmış örneklemede örnekleme çerçevesi kendi içlerinde homojen ve birbiriyle çakışmayan alt-gruplara (“tabaka” olarak adlandırılan) bölünür ve her bir alt-grup içerisinde birer basit rastgele örnek çekilir. Bin firmadan oluşan bir listeden 200 firmanın seçildiği hemen önceki misalimizde, büyüklüklerine dayanarak firmaları büyük (500 çalışandan fazla) , orta (50 ve 500 arasında çalışan) ve küçük (50 çalışandan az) olmak üzere kategorize ederek başlayabiliriz. Daha sonra, 200 firmalık örneğimizi teşkil etmek üzere, her bir alt-gruptan 67 firmayı rastgele seçebiliriz. Mamafih, örnekleme çerçevesinde büyük firmalardan çok daha fazla sayıda küçük firmalar mevcut olduğu için, eğer küçük, orta ve büyük firmaları eşit sayılarda örneklersek, seçilen örnek popülasyonu daha az temsil edecektir (yani, popülasyonda daha az sayıda yer alan büyük firmalar lehine yanlı olacaktır). Buna *orantısalmayın* tabakalı örnekleme adı verilir çünkü alt-grupların örnekteki orantıları örnekleme çerçevesindeki (veya ilgilenilen popülasyondaki) orantıları yansıtmaz ve daha küçük hacimdeki alt-gruplar (misalimizde büyük firmalar) *aşkın-örneklenmiş* olur. Alternatif bir teknik alt-grup örneklerini onların popülasyondaki hacimlerine orantılı olarak seçmektir. Mesela, eğer 100 büyük firma, 300 orta-büyük firma ve 600 küçük firma varsa, “büyükler” grubundan 20 firmayı, “orta-büyükler” grubundan 60 firmayı ve “küçükler” grubundan 120 firmayı örnekleyebiliriz. Bu halde, firmaların popülasyondaki orantısalm dağılımı örneğe de yansıtılmış olur ve buna istinaden bu teknik *orantısalm* tabakalı örnekleme olarak adlandırılır. Şuna dikkatini çekmek istiyorum ki orantısalm-olmayan yaklaşım mesela büyük firmalar gibi küçük hacimli alt-grupların temsilinde bilhassa efektiftir ve orantısalm yaklaşımla kıyaslandığında, illaki popülasyonu daha az temsil ediyor olması gerekmez, yeter ki orantısalm-olmayan yaklaşımın bulguları popülasyondaki alt-grupların orantılarına uygun olarak ağırlıklandırılabilir.

Öbek örnekleme. Eğer geniş bir coğrafi bölgeye yayılmış bir popülasyonunuz varsa, popülasyonun tamamının bir basit rastgele örneğini almak fizibil olmayabilir. Böyle bir durumda, popülasyonu “öbeklere” (genellikle coğrafi sınırlar ile) ayırmak, öbeklerin bir kaçını rastgele örnekleme ve bu öbeklerin her biri içindeki ünitelerin *tümünü* ölçmek akıllıca olabilir. Mesela, New York eyaletindeki yerel yönetimleri (muhtarlıklar diyebiliriz) örnekleme istiyorsak, yerel yöneticilerle görüşmek için, (bir basit rastgele örnekte yapmak zorunda kalabileceğimiz üzere) eyaleti boydan boya dolaşmak yerine, bulundukları bölgeleri baz alarak bu yerel yönetimleri öbeklendirebiliriz, üç bölgeden oluşan bir örnek seçebiliriz ve bu bölgelerin *her birindeki* yerel yöneticilerle görüşebiliriz. Mamafih, öbekler arasındaki farklılıklara bağlı olmakla beraber, bir öbek örneklemede örnek kestirimlerinin değişkenliği, bir basit rastgele örneğe göre, genel olarak daha yüksek olacaktır ve bundan ötürü sonuçlar, basit rastgele örnekten elde edilecek sonuçlara göre, popülasyona daha az genelleştirilebilir olacaktır.

Eşli-çiftler örnekleme. Bazen, araştırmacılar, bir spesifik kriteri baz almak suretiyle, bir popülasyon içindeki iki alt-grubu karşılaştırmak isteyebilirler. Mesela, niçin bazı firmalar diğerlerine göre istikrarlıca daha kârlıdır? Böyle bir öğrenme-çalışması düzenlemek için, firmaların bir örnekleme çerçevesini, total marjınlarını, hisse başına getirilerini veya bir başka kârlılık ölçüsünü baz alarak, “yüksek kârlı” firmalar ve “düşük kârlı” firmalar olarak kategorize etmemiz gerekir. Bundan sonra, bir alt-gruptaki firmaların bir basit rastgele örneğini alabiliriz ve seçtiğimiz firmaların her birini ikinci alt-gruptaki bir firmayla, büyüklüklerini, endüstri segmentlerini ve/veya başka bir eşleştirme kriterini baz alarak eşleştiririz. Şimdi, yüksek-kârlı ve düşük-kârlı firmaların daha detaylıca öğrenmeye-çalışabileceğimiz eşli iki örneğini elde etmiş olduk. Eşli-çiftler örnekleme tekniği belirli bir popülasyon içinde yer alan farklı gruplar arasındaki iki-kutuplu farkları anlamının ideal bir yoludur çoğu zaman.

Aşamalı örnekleme. Yukarıda betimlenen olasılıklı örnekleme tekniklerinin hepsi tek-aşamalı birer örnekleme tekniğidir. Örnekleme gereksinimlerimize bağlı olarak, bu tek-aşamalı teknikleri kombine etmek suretiyle aşamalı bir örnekleme yürütebiliriz. Mesela, firmaların bir

listesini, firma büyüklüğünü baz almak suretiyle, tabakalandırabiliriz ve sonra her bir tabaka içinde sistematik örnekleme yapabiliriz. Aynı minvalde, New York eyaletindeki okul bölgelerinin bir öbek örnekleme ile başlayabiliriz ve her bir öbek içinde okulların birer basit rastgele örneğini seçebiliriz ve her bir okul içinde öğrenim düzeylerinin birer basit rastgele örneğini seçebiliriz ve her bir öğrenim düzeyi içinde öğrencilerin birer basit rastgele örneğini öğrenme-çalışmamız için seçebiliriz. Bu durumda, dört-aşamalı bir örnekleme prosesimiz olur ki bu proses öbek örnekleme ve basit rastgele örnekleme ihtiva eder.

Olasılıksız Örnekleme

Olasılıksız örnekleme öyle bir örnekleme tekniğidir ki bu teknikle popülasyonun bazı ünitelerinin seçilme şansı *sıfırdır* veya ünitelerin seçilme şansı isabetlice belirlenebilir. Tipik olarak, üniteler, mesela kota veya elverişlilik gibi rastgele-olmayan kriterler baz alınarak seçilir. Seçim rastgele yapılmadığı için, olasılıksız örneklemede örnekleme hatalarının kestirimi mümkün olmaz ve dolayısıyla örnekleme yanlılığı kaçınılmaz olabilir. Bu yüzden, bir örneğin taşıdığı enformasyon popülasyona geri genelleştirilemez. Olasılıksız örnekleme tiplerinden bazıları şunlardır:

Elverişlilik örnekleme. Aynı zamanda tesadüf veya fırsat örnekleme olarak ta adlandırılan bu teknikte, bir örnek, popülasyonun araştırmacıya daha yakın olan, daha ulaşılabilir olan veya daha elverişli olan kısmından çekilir. Mesela, yaşadığın şehirdeki insanların bazı alış-veriş davranışlarını öğrenmeye-çalışmak için, eğer bir alış-veriş merkezinin önünde durursan ve yanından geçen kişilere hazırladığın yoklama anketini doldurtursan veya onlarla görüşme yaparsan, elde edeceğin yanıtlayıcılar örneği bir elverişlilik örneği olacaktır. Bu bir elverişlilik örneğidir çünkü başka alış-veriş merkezlerinde alış-veriş yapan tüm insanları sistematik olarak dışlamış oluyorsun. Seçmiş olduğun örnekte elde edeceğin görüşler mesela içindeki mağazaların mahiyeti (mesela lüks mağazalar daha varlıklı kesimi çeker), müşterilerinin demografik profili veya lokasyonu (mesela bir üniversiteye yakın bir alış-veriş merkezinin başlıca müşterileri o üniversitenin kendine has satınalma alışkanlıklarına sahip öğrencileridir) gibi bu alış-veriş merkezine münhasır karakteristikleri yansıtacaktır ve böylece geniş manada alış-veriş yapanlar popülasyonunun görüşlerinin temsilcisi olmayabilecektir. Bundan ötürü, böylesi gözlemlerin bilimsel genelleştirilebilirliği çok kısıtlı olacaktır. Elverişlilik örneklemesinin başka misalleri belirli bir sınıftaki öğrencileri örneklemek veya belirli bir tıp kliniğine gelen hastaları örneklemektir. Bu tip örnekleme pilot testler için yararlıdır ki bir plot testin amacı genelleştirilebilir çıkarımlardan daha ziyade enstrüman testi veya ölçüm geçerlemesidir.

Kota örnekleme. Bu teknikte, popülasyon karşılıklı-engelleyen alt-gruplara parçalanır ve önceden belirlenen bir kotaya uygun olarak her bir alt-gruptan rastgele-olmayan bir gözlemler kümesi seçilir. **Orantısal kota örneklemesinde**, her bir alt-gruptaki yanıtlayıcıların oranları popülasyonunkilerle eşleşmelidir. Mesela, eğer Amerikan popülasyonu %70 Kafkasyanlardan, %15 Hispaniklerden ve %13 Afrikan-Amerikanlardan meydana geliyorsa ve biz 98 kişilik bir örnekte bunların oy verme tercihlerini anlamayı arzu ediyorsak, bir alış-veriş merkezinin önünde durup oradan geçen insanlara oy verme tercihlerini sorabiliriz. Fakat Hispanik görünüşlü insanların yanıtları 15 i bulduğu (veya Afrikan-Amerikanların yanıtları 13 ü veya Kafkasyanların sayısı 69 u bulduğu) zaman artık bu alt-gruptan insanlara soru sormayı kesmeli ve örneğimizin etnik kompozisyonu genel Amerikan popülasyonunun kompozisyonu ile eşitlenene kadar diğer alt-grupları örneklemeye devam etmeliyiz. **Orantısız-olmayan kota örnekleme** daha az kısıtlayıcıdır çünkü orantısız bir temsili gerektirmez fakat her grupta minimum bir hacmi doldurmak gerekir. Bu halde, her bir etnik alt-gruptan (Kafkasyanlar, Hispanikler ve Afrikan-Amerikanlar) 50 yanıtlayıcıyı dahil etmeye karar verebiliriz ve tüm

grupların kotalarını doldurduğumuzda durabiliriz. Kota örneklemesinin hiç bir tipi Amerikan popülasyonunun temsilcisi olmayacaktır çünkü öğrenme-çalışmamızın New York'ta veya Kansas'ta bir alış-veriş merkezinde yürütölüp yürütölmediğine bağılı olarak, sonuçlarımız tamamen farklı olabilecektir. Orantısal-olmayan teknik popülasyonun daha az temsilcisi olmakla beraber, küçük ve eksik-temsil-edilen grupların görüşlerinin yakalanmasına olanak sağladığı için yararlı olabilir.

Uzman örnekleme. Bu teknikle yanıtlayıcılar, öğrenilmeye-çalışılan fenomene dair uzmanlıkları baz alınmak suretiyle rastgele-olmayan bir yolla seçilirler. Mesela Sarbanes-Oxley yasası gibi hükümet politikalarının çarpan etkilerini (*impacts*) anlamak için, bu yasaya aşına olan tüzel saymanların bir grubu örneklenebilir. Bu yaklaşımın avantajı şudur ki uzman olmayanlara göre, uzmanlar öz-konuya daha aşına olmak durumunda olduğı için, uzmaların bir örneğinden elde edilecek görüşler hem uzmanları hem de uzman olmayanları içeren bir örneğe göre daha muteber olacaktır fakat yine de, elde edilen bulgular hala geniş manada popülasyonun tümüne genelleştirilebilir değildir.

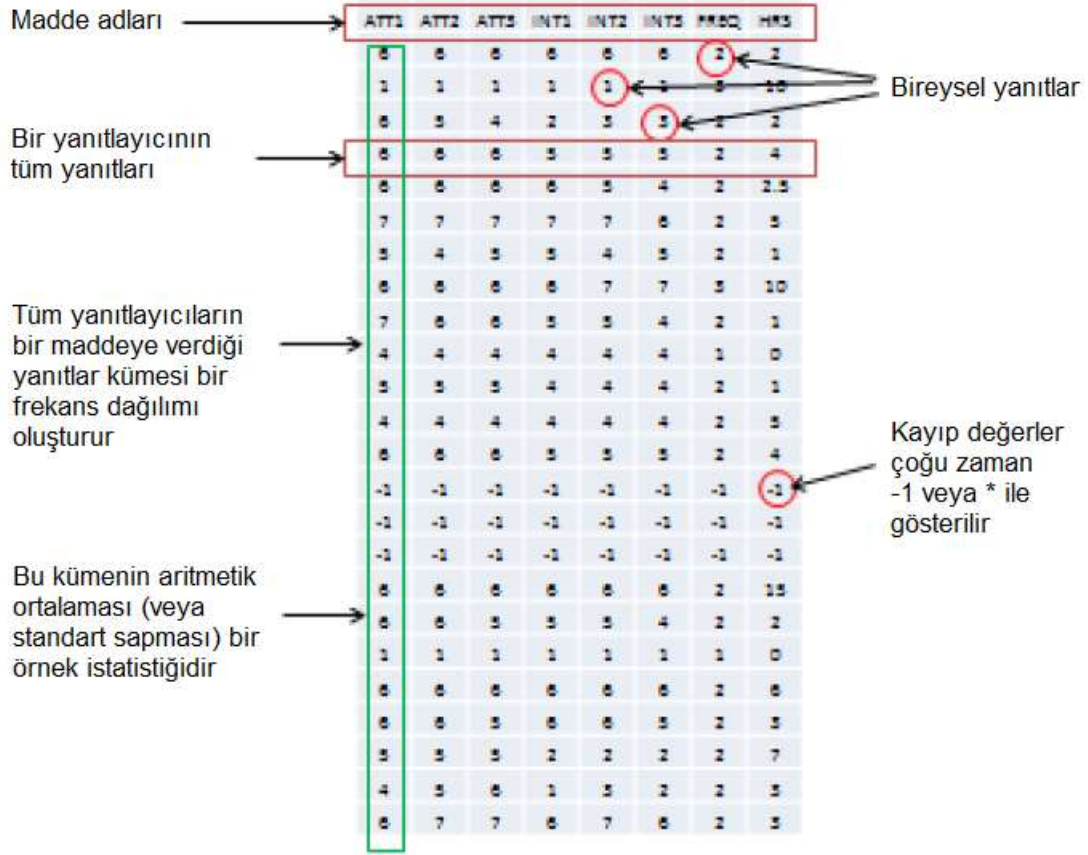
Kartopu örnekleme. Kartopu örneklemede, öğrenme-çalışmamıza dahil etme kriterimizle eşleşen bir kaç yanıtlayıcıyı teşhis etmekle başlarız ve daha sonra onlardan seçim kriterimize uyan başkla tanıdıklarını tavsiye etmelerini isteriz. Mesela, eğer kompütür ağı yöneticilerini yoklamayı arzu ediyorsak ve böyle yalnızca bir iki kişiyi tanıyorsak, onlarla başlayabiliriz ve daha sonra onlardan ağı yöneticiliğı yapan başkalarını bize tavsiye etmelerini isteyebiliriz. Bu metot, her ne kadar, soruşturmacıyı temsilci örneklere götürmese de, ulaşılması zor olan popülasyonlara ulaşmakta veya örnekleme çerçevesinin mevcut olmadığı zamanlarda tek çıkar yol olabilir.

Örnek İstatistikleri

Önceki kısımlarda, popülasyon parametresi, örnek istatistiğı ve örnekleme yanlılığı gibi terimlerle tanışmış olduk. Bu kısımda, bu terimlerin ne manaya geldiğini ve birbirleriyle nasıl ilişkili olduklarını anlamaya çalışacağız.

Mesela bir kişinin bir maddeye Likert-ölçeğindeki yanıtı gibi, belirli bir üniteye belli bir gözlemi ölçtüğümüz zaman, o gözleme yanıt adı verilir (Figür 8.2 ye bak). Bir **yanıt**, başka bir deyişle, örneklenen bir üniteye elde edilen ölçüm değeridir. Bir enstrümandaki farklı maddeler için, her yanıtlayıcı bize farklı yanıtlar verecektir. Farklı yanıtlayıcıların aynı maddeye veya gözleme ait yanıtları, vuku bulma frekanslarına istinaden bir **frekans dağılımı** grafiğı ile ifade edilebilir. Bir örnekteki çok sayıda yanıt için, bu frekans dağılımı **normal dağılım** adı verilen çan şeklinde bir eğriye benzeme eğilimi gösterir ki bu eğri örneğin bütününe ait mesela örnek aritmetik ortalaması (bir örnekteki tüm gözlemlerin ortalaması) veya standart sapması (bir örnekteki gözlemlerin değişkenliğı veya yayılımı) gibi sayısal betimleyicilerle karakterize edilir. Örneğe ait bu sayısal betimleyiciler **örnek istatistikleri** olarak adlandırılır (bir “istatistik” bir değerdir ki gözlenen veriden hesaplanır). Popülasyonlar da, keza, aritmetik ortalamalara ve standart sapmalara sahiplerdir öyle ki bunları elde edebilmek için popülasyonun bütününe örnekleme gerekir. Mamafih, çoğı zaman, popülasyonun bütünü asla gözlenemeyeceğı (veya örnekleneemeyeceğı) için popülasyona ait sayısal betimleyiciler çoğı zaman meçhuldür ve bunlara **popülasyon parametreleri** adı verilir. Bir rastgele örnekte veri toplamamızın sebebi işte bu popülasyon parametrelerinin birer *nokta kestirimi* olarak örnek istatistiklerini elde etmek istememizdir. Örnek istatistikleri popülasyon parametrelerinden farklı olabilirler eğer örnek popülasyonu mükemmelce temsil etmiyorsa. Bu ikisi (yani, istatistik ve parametre) arasındaki farka **örnekleme hatası** adı verilir. Teorik olarak, eğer örnek hacmini, örneğı popülasyona daha-ve-daha yaklaştırmak suretiyle, tedricen artırırsak, o zaman örnekleme hatası azalacak ve bir örnek istatistiğı tekabül ettiğı popülasyon parametresine daha-

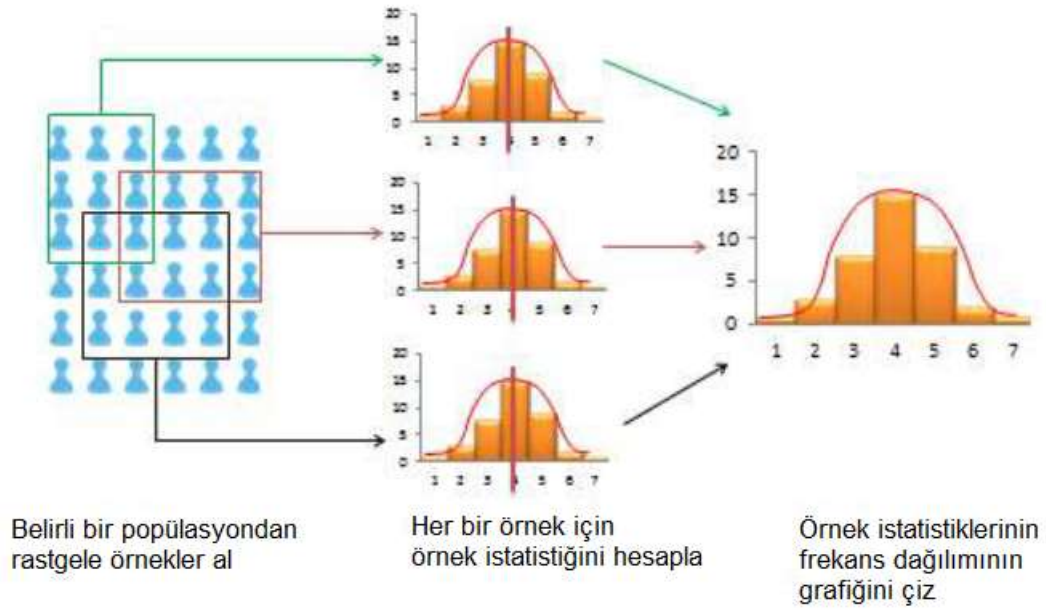
ve-daha yaklaşacaktır.



Figür 8.2. Örnek istatistiği

Eğer bir örnek içinden çekilmiş olduğu popülasyonun bihakkın temsilcisi ise, o zaman hesaplanan örnek istatistikleri tekabül ettikleri *meçhul* popülasyon parametreleri ile özdeş olacaklardır. Peki, örnek istatistiklerinin popülasyon parametrelerine en azından makulce yakın olup olmadıklarını nasıl bileceğiz? Burada, **örnekleme dağılımı** kavramını anlamaya ihtiyacımız var. Hayal et ki belirli bir popülasyondan, Figür 8.3 te gösterildiği üzere, üç farklı rastgele örnek aldın ve her örnek için mesela örnek aritmetik ortalaması ve standart sapması gibi istatistikleri türettin. Eğer her bir rastgele örnek popülasyonun bihakkın temsilcisi ise, o zaman üç örneğin üç örnek aritmetik ortalaması özdeş (ve popülasyon parametresine eşit) olacaktır ve örnek aritmetik ortalamalarının değişkenliği sıfır olacaktır. Fakat bu aşırıca olabirliksizdir çünkü rastgele örneklerin her biri popülasyonun farklı bir alt kümesini oluşturabilecek ve bundan ötürü bunların aritmetik ortalamaları birbirlerinden hafifçe farklı olabilecektir. Mamafih, bu üç örnek aritmetik ortalamasını alarak, örnek aritmetik ortalamalarının bir histogramını çizebiliriz. (**Histogram** bir frekans dağılımının bir grafiğidir.) Bu örneklerin, eğer sayısını üçten 10 a, 100 e veya 1000 e çıkarırsak, histogram bir örnekleme dağılımı olur. Bundan ötürü, bir örnekleme dağılımı *aynı hacimdeki örnekler kümesinden* gelen (örnek aritmetik ortalaması gibi) bir *örnek istatistiğinin* frekans dağılımıdır. Öte yandan, frekans dağılımı denildiği zaman kastedilen *tek bir örnekten* gelen bir *yanıtın* (gözlemin) frekans dağılımıdır. Tam olarak bir frekans dağılımında olduğu gibi, hakeza bir örnekleme dağılımında da, örnek istatistiklerinin daha çoğu kendi aritmetik ortalamaları (ki bu da popülasyon parametresinin bir kestirimi olarak kabul edilir) etrafında öbikleşme ve daha azı kendi aritmetik ortalamaları etrafında saçılma temayülü gösterir. Örneklerin sayısı sonsuza giderken, bu dağılımın bir normal dağılıma yaklaştığı varsayılır. Bir örnek istatistiğinin örnekleme dağılımındaki değişkenliğine veya yayılımına (yani, bir örnek istatistiğinin standart

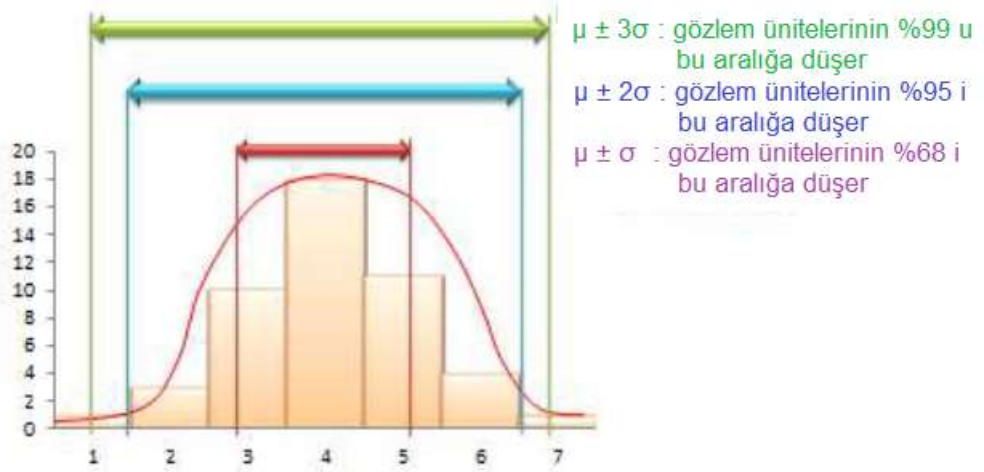
sapmasına) **standart hata** adı verilir. Öte yandan, standart sapma terimi tek bir örnekte gözlenen yanıtın değişkenliğini ifade etmek için kullanılır.



Figür 8.3. Örnekleme dağılımı

Bir örnek istatistiğinin örnekleme dağılımının aritmetik ortalamasının değerinin teorik olarak meçhul popülasyon parametresine eşit olduğu varsayılır. Örnekleme dağılımının yayılımına (yani, standart hatasına) istinaden, bununla birlikte, popülasyon parametresinin kestirimi için bir *güven aralığı* tesis etmek de olanaklıdır, aynı zamanda. **Güven aralığı** bir örnek istatistiği etrafında popülasyon parametresini belirli bir olasılıkla içinde bulunduran spesifik bir aralıktır. Tüm normal dağılımlar 68-95-99 yüzde kuralına (Figür 8.4 e bak) uyma eğilimindedir ki bu kural dağılımdaki gözlemlerin yaklaşık %68 inin aritmetik ortalamasının 1 standart sapma civarında ($\mu \pm 1\sigma$), yaklaşık % 95 inin aritmetik ortalamasının 2 standart sapma civarında ($\mu \pm 2\sigma$) ve yaklaşık %99 unun aritmetik ortalamasının 3 standart sapma civarında ($\mu \pm 3\sigma$) yer aldığını söyler. Bir örnekleme dağılımının, örneklerin sayısı sonsuza giderken, teorik olarak normal dağılıma yaklaşacağı varsayıldığı için, 68-95-99 kuralı uygulanmak suretiyle, söylenebilir ki:

- (Örnek istatistiği ± 1 standart hata) popülasyon parametresi için %68 lik bir güven aralığını temsil eder,
- (Örnek istatistiği ± 2 standart hata) popülasyon parametresi için %95 lik bir güven aralığını temsil eder,
- (Örnek istatistiği ± 3 standart hata) popülasyon parametresi için %99 luk bir güven aralığını temsil eder.



Figür 8.4. Güven aralığı için 68-95-99 yüzde kuralı

Eğer bir örnek “yanlı” ise (yani, popülasyonun temsilcisi değilse), o örnekten hesaplanan bir istatistiğin örnekleme dağılımının 68-95-99 kuralına uyması beklenmez. Mesela, ileride yapabileceğin bir regresyon analizinde, örnek istatistiğinin (yani, örnek regresyon katsayısının) tekabül ettiği popülasyon parametresi için %95 lik bir güven aralığı kestiriminin isabetliliği yapacağın örnekleme yanlı olup olmamasına bağlıdır. Güven aralıklarının kestirilmesinin yanısıra, popülasyon parametrelerinin değerleri üzerine kurulan hipotezlerin testi de yine örnek istatistiklerinin örnekleme dağılımlarına istinaden yapılmaktadır. Mesela bir kalite kontrol prosedürünün parçası olarak, bir üretim bandından alınan bir örneğin kusurlu parça orantısının anlamlılığı %0.1 düzeyinde test edilebilir.

Bölüm 9

Yoklama Araştırması

Yoklama araştırması (/survey research) insanlar ve onların tercihleri, düşünceleri ve davranışları hakkında sistematik olarak veri toplamak için standartlaştırılmış anketlerin (/questionnaires) veya görüşmelerin (/interviews) kullanılmasını ihtiva eden bir araştırma metodudur. Her ne kadar nüfus sayımı yoklamalarının yürütülmesinin tarihi Antik Mısır'a kadar uzanıyor olsa bile, formal bir araştırma metodu olarak sosyolog Paul Lazarsfield tarafından Amerika Birleşik Devletleri'nde politik görüşlerin oluşmasına radyonun etkilerini incelemek için 1930-40'larda öncülüğü yapılmıştır. O zamandan beri, bu metot sosyal bilimlerde nicel araştırma için çok popüler bir metot olmuştur.

Yoklama metodu betimsel, keşifsel veya açıklayıcı araştırma için kullanılabilir. Bu metot analiz ünitesi olarak bireysel insanların yer aldığı öğrenme-çalışmaları için en uygundur. Mesela gruplar, organizasyonlar veya dyadlar (mesela alıcılar ve satıcılar gibi organizasyon çiftleri) gibi başka analiz üniteleri de keza yoklamalar kullanılarak öğrenilmeye-çalışılsa bile, böylesi öğrenme-çalışmaları çoğu zaman her bir üniteden spesifik bir kişiyi o ünite için bir "kilit enformasyon-temin-edici" veya "vekil" olarak kullanır ve böylesi yoklamalar yanıtlayıcı yanlılığına açık olabilir eğer seçilen enformasyon-temin-edici kifayetli bilgiye sahip değilse veya ilgilenilen fenomene dair yanlı bir görüşe sahipse. Mesela, CEO'lar kendi şirketlerindeki çalışanların algılarını veya takım-çalışmasını kifayetle bilmeyebilirler ve bu yüzden çalışanların özgüvenini veya takım dinamiklerini öğrenme-çalışmalarında yanlı enformasyon-temin-edici olabilirler.

Yoklama araştırması, başka araştırma metotlarına kıyasla bazı güçlü içkin hususiyetlere sahiptir. Birincisi, yoklamalar mesela insanların tercihleri (diyelim politik eğilim), vasıfları (diyelim özgüven), tutumları (diyelim göçmenlere karşı), inançları (diyelim büyüye dair), görüşleri (diyelim yeni bir yasa hakkında), davranışları (diyelim tütün veya alkol içme) veya olgusal enformasyonları (diyelim gelir) gibi gözlenemeyen çok çeşitli verinin ölçümü için mükemmel bir araçtır. İkincisi, yoklama araştırması aynı zamanda direk olarak gözlenemeyecek kadar geniş alana yayılmış bir popülasyon hakkında uzaktan veri toplamanın ideal bir yoludur. Mesela bir ülkenin tamamı gibi geniş bir alanı posta, e-posta veya telefon yoklamaları kullanmak suretiyle kapsayabiliriz ve istihdam ettiğimiz örnekleme prosedürü ile popülasyonun küçük bir örnekte kifayetle temsil edildiğinden emin olabiliriz. Üçüncüsü, yanıtlayıcılara musallat olunmaması ve insanların kendilerinin uygun gördükleri yanıtları verebilmelerine olanak sağlaması nedeniyle anket yoklamaları bazı yanıtlayıcılarca tercih edilir. Dördüncüsü, her hangi bir örnekleme çerçevesinin mevcut olmadığı mesela evsizler veya yasadışı göçmenler gibi belli popülasyon gruplarına ulaşmanın belki de yegane yolu görüşmeler olabilir. Beşincisi, geniş örnek yoklamaları, çoklu değişkenlerin analizinde bile, küçük etkilerin saptanmasına olanak sağlayabilir ve yoklama dizaynına bağlı olarak, popülasyon alt-gruplarının karşılaştırmalı analizine (yani, grup-ici ve gruplar-arası analiz) olanak sağlayabilir. Altıncısı, yoklama araştırması araştırma süresi, efor ve maliyet bakımından mesela deneysel araştırma ve vaka araştırması gibi başka metotlara göre daha ekonomiktir. Aynı zamanda, yoklama araştırması kendine mahsus bazı dezavantajlara da sahiptir. Mesela yanıt-yok yanlılığı, örnekleme yanlılığı, sosyal beğenilirlik yanlılığı ve anımsama yanlılığı gibi pek çok yanlılığa açıktır ki bunlar bu bölümün son kısmında tartışılacaktır. Verinin nasıl toplanacağına bağlı olarak, yoklama araştırması iki geniş kategoriye ayrılabilir: anket yoklamaları (ki bunlar posta, gruplu veya çevrimiçi yoklamalar olabilir) ve görüşme yoklamaları (ki bunlar yüz-yüze,

telefon veya odak grup görüşmeleri olabilir). Anketler yanıtlayıcıların yazarak doldurdıkları enstrümanlarken, görüşmeler yanıtlayıcılar tarafından temin edilen sözlü yanıtlara istinaden görüşmeci tarafından doldurulur. Aşağıda tartışıldığı üzere, her tipin maliyeti, hedef popülasyonu kapsayıcılığı ve soruları sorarken, araştırmacıya sağladığı esneklik bakımından kendince güçlü ve zayıf tarafları vardır.

Anket Yoklamaları

Anket öyle bir araştırma enstrümanıdır ki yanıtlayıcıların yanıtlarını standartlaştırılmış bir şekilde almak maksadıyla hazırlanan bir sorular (maddeler) kümesinden oluşur. Sir Francis Galton (1822-1911) tarafından icat edilmiştir. Sorular sütüktürsüz (*unstructured*) veya sütüktürlü (*structured*) olabilir. Sütüktürsüz sorulara, yanıtlayıcıdan kendi kelimeleriyle bir yanıt vermesi istenirken, sütüktürlü sorular için, yanıtlayıcıdan verilen şıklar arasından birisini seçmesi istenir. Sütüktürlü bir ankette, deneklerin bireysel sorulara (maddelere) verdikleri yanıtlar, istatistiksel analiz için, kompozit bir ölçekte veya endekte toplaştırılabilir. Sorular öyle dizayn edilmelilerdir ki yanıtlayıcılar onları manalı bir şekilde okuyabilsinler, anlayabilsinler ve yanıtlayabilsinler. Bu açıdan bakıldığında, yoklama metodu mesela çocuklar veya hiç öğretim almayanlar gibi demografik gruplar için mütenasip veya pratik olmayabilir.

Anket yoklamalarının çoğu kendi-kendine uygulanan posta yoklamaları olarak yürütülür ki bunlarda aynı anket formu ön-ödemeli bir geri-gönderim zarfı ile beraber çok sayıda insana postalanır ve istekli yanıtlayıcılar yoklamayı canları istediği gibi doldurup geri postalarlar. Posta yoklamaları şu bakımdan avantajlıdır ki yanıtlayıcılara musallat olunmaz ve pek çok ülkede yığın-posta ucuz olduğu için uygulanmaları pahalı değildir. Mamafih, posta yoklamalarının yanıt oranı hayli düşüktür çünkü çoğu insan yoklama talebini görmezden gelmeyi yeğler. Aynı zamanda, yanıtlayıcıların anket formlarını doldurmalarında ve geri postalamalarında uzun gecikmeler (bir kaç ayı bulan) olabilir. Bundan ötürü, araştırmacı yanıtların geri dönüşlerini sürekli göz altına almalı ve yanıtlarını göndermeyenleri izlemeli ve anımsatmalarda bulunmalıdır (bir veya 1.5 aylık aralıklarla iki veya üç anımsatma idealdir). Anket yoklamaları keza yanıtlayıcıya açıklama yapılmasını gerektiren veya detaylı yazılı yanıt gerektiren meseleler için pek te uygun değildir. Boylamsal (*longitudinal*) dizaynlar aynı yanıtlayıcıların bir kümesini farklı zamanlarda yoklamak için kullanılabilir fakat yanıt oranları bir yoklamadan diğerine hızla düşme eğilimi gösterir.

İkinci bir yoklama tipi **gruplu ankettir**. Müşterek bir yerde ve zamanda, yanıtlayıcıların bir örneği biraraya getirilir ve her bir yanıtlayıcıdan yoklama anketini o mekandayken doldurmaları istenir. Yanıtlayıcılar yanıtlarını birbirleriyle etkileşmeden bağımsızca girerler. Bu format araştırmacı için elverişlidir ve yüksek yanıt oranından emin olunur. Eğer yanıtlayıcıların anlayamadığı spesifik sorular olursa, açıklama isteyebilirler. Çoğu organizasyonda, çalışanların bir grubunu bir konferans veya yemek salonunda biraraya getirmek nispeten kolaydır, bilhassa eğer yoklama yönetimin onayı ile yapılıyor ise.

Son zamanlardaki bir anket yoklaması tipi çevrimiçi web yoklamasıdır. Bu yoklamalar interaktif formlar kullanılmak suretiyle internet üzerinden yürütülür. Yanıtlayıcılara, yoklamaya katılmalarını rica etmek için, yoklamanın doldurulabileceği bir çevrimiçi web sitesinin linkinin yer aldığı bir elektronik posta gönderilebilir. Alternatif olarak, yoklama bir e-postaya gömülebilir ve doldurulduktan sonra e-posta ile geri gönderilebilir. Bu yoklamaların yapımları hayli pahasızdır, sonuçlar anında bir çevrimiçi veritabanına kaydedilir ve yoklama eğer gerekirse kolayca modifiye edilebilir. Mamafih, yoklama web sitesi eğer şifre-korumalı değilse veya çoklu yanıt girişlerini engellemek üzere dizayn edilmemişse, yanıtlar kolayca zaafa uğrayabilir. Dahası, örnekleme yanlılığı anlamlı bir mesele olabilir çünkü yoklama mesela fakir, yaşlı veya azınlık grupları gibi kompütörü veya internet erişimi olmayan insanlara

ulaşabilmez ve yanıtlayıcılar örneği zamanlarının çoğunu çevrimiçi geçiren ve böylesi yoklamaları doldurmaya zamanı ve yeteneği olan daha genç demografik segmente doğru çarpılır. Yanıt oranını hesaplamak problemlili olabilir eğer yoklama linki yalnızca e-posta ile doğrudan hedeflenen yanıtlayıcılara gönderilmek yerine, başka sosyal platformlarda da paylaşılsa. Bu sebeplerden ötürü, çoğu araştırmacı çifte-medya yoklamalarını (mesela, posta yoklaması ve çevrimiçi yoklama) tercih etmektedir ki yanıtlayıcılar kendi beğendikleri yanıtlama metodunu seçebilsinler.

Yoklama anketi inşa etmek bir sanattır. Soruların içeriği, kelimelere dökülmeleri, formatları ve sıralamaları hakkında pek çok karar vermeyi gerektirir ve bunların her biri yoklamanın yanıtları için önemlidir.

Yanıt formatları. Yoklama soruları sütüktürlü veya sütüktürsüz olabilirler. Sütüktürlü soruların yanıtları şu yanıt formatlarından biri ile alınır:

- *Dikotomik yanıt*; öyle ki yanıtlayıcılardan mesela doğru/yanlış, evet/hayır veya katılıyorum/katılmıyorum gibi olanaklı iki seçenekten birini seçmeleri istenir.
- *Adlandırılmış (/nominal) yanıt*; öyle ki yanıtlayıcılara mesela “hangi sektörde istihdam ediliyorsun? İmalat / tüketici hizmetleri / perakende / öğretim / sağlık / turizm / başka” gibi iki veya daha çok sayıda sırasız (/unordered) opsiyon sunulur.
- *Sıralı (/ordinal) yanıt*; öyle ki yanıtlayıcılara mesela “En yüksek öğrenim düzeyiniz nedir? İlk/orta/lise/önlisans/lisans/lisansüstü” gibi iki veya daha çok sayıda sıralı (/ordered) opsiyon sunulur.
- *Aralık-düzeyinde yanıt*; öyle ki yanıtlayıcılara eşit mesafede sıralı seçenekler sunulur. Bu seçeneklerin sayıyla ifade edilebilir olması gerekir. Beş-puanlı veya 7-puanlı Likert ölçeği, semantik diferansiyel ölçek veya Guttman ölçeği ile elde edilen yanıtların, sosyal araştırmada, her ne kadar, aralık düzeyinde olduğu farzedilse bile, esasında sıralı yanıtlardır.
- *Sürekli yanıt*; öyle ki yanıtlayıcılar yanıtlarını eşit aralıklı ve manalı bir sıfır noktasına sahip bir ölçek üzerinden sürekli bir değer olarak genellikle boş bırakılmış bir alana girerler. Mesela yanıtlayıcıların yaşları, firmadaki çalışma süreleri veya aylık kazançları gibi.

Soru içeriği ve lafız (/wording). Yoklama araştırmasında alınan yanıtlar sorulan soruların tiplerine karşı çok duyarlıdır. Kötü ifade edilmiş veya muğlak soruların pek te değeri olmayan manasız yanıtlar ile sonuçlanacağını söyleyebiliriz. İyi yoklama soruları yaratmak için, Dillman (1978) her bir sorunun aşağıdaki hususiyetleri bakımından dikkatlice incelenmesini tavsiye etmiştir:

- *Soru açık ve anlaşılabilir mi?* Yoklama soruları çok basit bir dille ifade edilmelidir. Aktif cümleler tercih edilmeli ve bazı yanıtlayıcılar tarafından anlaşılamayacak komplike kelimelerden veya jargondan kaçınılmalıdır. Anketin tüm soruları benzer bir tarzda lafzedilmelidir ki yanıtlayıcılar onları daha kolay okuyabilsinler ve anlayabilsinler. Eğer hedef popülasyonumuz mesela doktorlar, avukatlar veya akademisyenler gibi kendi günlük ortamlarında zaten bir jargonla konuşan özel bir grup ise, bu istisnai bir durumdur elbette.
- *Soru negatif bir tavırla mı lafzediliyor?* Mesela “Yerel yönetim vergileri artırmamalı mı?” gibi negatif bir tavırla lafzedilmiş sorular yanıtlayıcıların çoğunun kafasını karıştırmaya ve isabetsiz yanıtlara yol açmaya meyillidir. Böyle sorulardan kaçınmak gerekir ve herhalükârda çifte-negatiflerden uzak durmak gerekir.
- *Soru muğlak mı?* Yoklama soruları öyle lafzedilmeli veya ifade edilmelidir ki farklı yanıtlayıcılar tarafından farklıca yorumları yapılabilmemelidir. Mesela, eğer

yanıtlayıcıya “Yıllık gelirin nedir?” diye sorarsan, maaşını mı, hisselerini mi, kiralarını mı veya başka gelirini mi; kişisel gelirini mi yoksa aile gelirini mi kastettiğin açık değildir. Farklı yanıtlayıcıların farklı interpretasyonları karşılaştırılabilir yanıtla yol açar ve bu yanıtların interpretasyonu da doğru olarak yapılabilmektedir.

- *Sorunun lafzı yanlı mı veya değer-empoze-edici mi?* Yanlılık bir sorunun denekleri belli bir yönde cevap vermeye teşvik eden herhangi bir özelliğine atıfta bulunur. Kenneth Rasinky (1989) halkın hükümet harcamalarına karşı tutumu üzerine yapılmış çok sayıda öğrenme-çalışmasını inceledi ve gözledi ki yanıtlayıcılar “fakirlere yardım” lehine daha fazla destek vermeye ve “refah” lehine daha az destek vermeye eğilimlidir, her ikisi de aynı manaya geliyor olmasına rağmen. Bu öğrenme-çalışmasında, aynı zamanda, “yükselen suç oranını durdurma” lehine daha fazla destek (ve “yasal yaptırımlar” lehinde daha az), “büyük şehirlerin problemlerini çözme” lehine daha fazla destek (ve “büyük şehirlere yardım” lehine daha az) ve “uyuşturucu madde bağımlılığı ile mücadele” lehine daha fazla destek (ve “uyuşturucu madde rehabilitasyonu” lehine daha az) gözlemlendi. Yanlı bir dil veya tonlama gözlenen yanıtların çarpılmasına yol açar. Yanlı bir lafzı önceden farkedebilmek çoğu zaman zordur, fakat yanlı bir dilden kaçınmak için, yoklama soruları azami dikkatle incelenmelidir.
- *Soru çift-namlulu mu?* Çift-namlulu sorular çoklu cevaplar verilebilen sorulardır. Mesela, işin için tahsis edilen donanımdan ve yazılımdan memnun musun? Bu misalde, eğer yanıtlayıcı donanımdan memnun fakat yazılımdan memnun değilse veya tersiye, nasıl bir cevap verecek? Çift-namlulu soruları birbirinden ayrı sorulara ayırmak daima tavsiye edilir, şöyle ki: (1) “İşin için tahsis edilen donanımdan memnun musun?” ve (2) “İşin için tahsis edilen yazılımdan memnun musun?” Başka bir misal: Aileniz TRT1 TV kanalını beğeniyor mu? Ailenin bazı üyeleri bazı program türlerini beğeniyor veya beğenmiyor olabilirler fakat hepsi hepsini değil çoğu zaman.
- *Soru fazlaca genel mi?* Bazen, fazlaca genel sorular yanıtlayıcının algısını isabetlice nakledebilmez. Eğer birisine belli bir kitaptan ne kadar hoşlandığını sorarsan ve “hiç”ten “fazlaca”ya giden bir yanıt ölçeği verirsen ve eğer o kişi “fazlaca”yı seçerse, o kişi ne kastetmiş olur? Onun yerine, mesela “bu kitabı başkalarına tavsiye edecek misin?” veya “Aynı yazarın başka kitaplarını okumayı düşünüyor musun?” gibi daha spesifik davranışsal sorular sorulması daha yararlıdır. Aynı minvalde, “Firmanızın büyüklüğü nedir?” sorusu yerine firmanın büyüklüğünü ölçen “Firmanızda kaç eleman çalıştırıyorsunuz?” ve/veya “Firmanızın yıllık gelir nedir?” sorularını sormak daha yararlıdır.
- *Soru fazlaca detaylı mı?* Araştırmanın maksadına hiç bir şekilde hizmet etmeyen fazlaca detaylı sorulardan kaçınmalısın. Mesela, bir hanedeki her çocuğun yaşına mı ihtiyacın var yoksa hanedeki çocuk sayısı işini görür mü? Mamafih, eğer emin olamıyorsan, daha detaylı daha genelden daha iyidir.
- *Soru faraziyeli mi?* Eğer “Vergi indirimini nasıl faydalı buluyorsunuz?” diye sorarsak, yanıtlayıcının vergi indirimini faydalı bulduğunu farzetmiş oluruz. Oysa, pek çok insan vergi indirimlerini faydalı bulmayabilir çünkü vergi indirimleri genellikle devlet okullarına daha az fon ayrılmasına, daha kalabalık sınıflara ve polis, ambulans ve itfaiye gibi hizmetlerin daha azaltılmasına yol açar. Kendi içinde faraziyeler barındıran sorulardan kaçınmalıyız.
- *Soru hayal mahsulü mü?* Pek çok TV yarışma programlarının popüler sorusu şudur: “Eğer bu yarışmada bir milyon \$ kazanırsan, nasıl harcamayı planlıyorsun?” Yanıtlayıcıların pek çoğu hayatları boyunca bu miktarda bir parayla karşılaşmadıkları

gibi hakkında hiç düşünmemişlerdir bile (pek çoğu hatta vergi kesintilerinden sonra Amerika Birleşik Devletleri'nde ellerine yaklaşık 640 bin \$ geçeceğini ve ödemenin yirmi yıla kadar bir periyoda yayılacağını dahi bilmez) ve bundan ötürü yanıtlayıcıların verecekleri yanıtlar mesela dünya turuna çıkmak, bir restoran veya bar satın almak, öğrenime harcamak, emeklilik için saklamak veya ihtiyaç sahiplerine yardım etmek gibi hayli rastgele olacaktır. Hayali soruların hayali cevapları olur ki bilimsel çıkarımlarda bulunmak için kullanılabilir.

- *Yanıtlayıcılar soruyu cevaplamak için gereken enformasyona sahip mi?* Çoğu zaman, deneklerin bir soruyu cevaplamak için gereken enformasyona sahip olduklarını varsayarız oysa gerçekte sahip olmadıkları halde. Böyle bir durumda, yanıt alınmış olsa dahi, sorulan soru hakkındaki bilgi eksikliğinden ötürü yanıtlar isabetsiz olmaya meyillidir. Mesela, bir şirketin CEO'suna bilgisi dahilinde olmayan günlük operasyonlar hakkında bir soru veya öğretmenlere öğrencilerinin ne kadar öğrendiğini veya lise öğrencilerine “ABD hükümetinin Domuzlar Körfezi krizinde mütenasip davrandığını düşünüyor musun?” sorusunu sormamalıyız.

Soru dizilimi. Genelde, sorular birinden diğerine doğru mantıksal olarak akmalıdır. En iyi yanıt oranlarını yakalamak için, sorular en az duyarlı olandan en fazla duyarlı olana doğru, olgusal veya davranışsal olandan tutumsal olana doğru ve daha genel olandan daha spesifik olana doğru akmalıdır. Soru dizilimi için bazı genel kurallar şunlardır:

- Çabuk cevaplanabilecek kolay ve göz-korkutmayan sorularla başla. Birey-bazlı yoklamalarda mesela yaş, cinsiyet ve öğrenim düzeyi gibi demografik sorular ve organizasyon-düzeyinde yoklamalarda mesela çalışan sayısı, yıllık gelir veya endüstri kolu gibi firmografik sorular iyi opsiyonlardır.
- Asla açık uçlu bir soruyla başlama.
- Eğer olayların tarihsel bir sırasını takibetmek icabediyorsa, en öncekinden en sonrakine doğru kronolojik bir sıra takibet.
- Bir defada bir mevzuyu sor. Mevzular arasında gidip-gelmek icabediyorsa, mesela “Sonraki kısım sizin ... hakkındaki görüşlerinizi inceliyor” gibi bir geçiş kullan.
- Gerekli yerde, mesela “Eğer 5 inci soruya “evet” cevabını vererseniz, lütfen 2 inci kısma ilerleyin. Eğer “hayır” cevabını vererseniz 3 üncü kısma ilerleyin.” gibi filtre veya kontenjans soruları kullan.

Diğer altın kurallar. Sana nasıl davranılmasını arzu ediyorsan, yanıtlayıcılara da öyle davran. Yanıtlayıcıların vaktine, katılımına, güvenine ve kişisel enformasyonun mahremiyetine özen göster. Tüm yoklama araştırmalarında şu stratejileri daima uygula:

- İnsanların vakti kıymetlidir. Onların vakitlerine saygılı ol. Yoklamayı mümkün mertebe kısa tut ve gereken ne ise onunla kısıtla. Yanıtlayıcılar herhangi bir yoklama için 10-15 dakikadan daha fazla zaman harcamaktan hoşlanmazlar, ne kadar önemli olduğuna bakmaksızın. Yoklama ne kadar uzarsa, yanıt oranı dramatik olarak düşer.
- Daima yanıtlayıcılara yanıtlarının mahremiyeti ve onlara ait verinin nasıl kullanılacağı ve sonuçların nasıl rapor edileceği hakkında güven ver.
- Organizasyonel yoklamalar için, yanıtlayıcılara nihai sonuçların bir kopyasını göndereceğine dair güven ver.
- Öğrenme-çalışmana katıldıkları için, yanıtlayıcılarına teşekkür et.
- Son olarak, anketini sahada fiilen uygulamadan önce, en azından bir elverişlilik örneği kullanarak bir ön-testini yap. Böyle bir ön-test soru lafzındaki muğlaklığı, açıklık eksikliğini veya yanlışlığı ifşa edebilir ki niyet edilen örneğe uygulanmadan önce bunlar elemine edilebilsin.

Görüşme Yoklaması

Görüşmeler anketlere göre daha kişiselleştirilmiş bir veri toplama metodudur ve anket yoklaması ile aynı araştırma protokolü (yani, standartlaştırılmış bir soru kümesi) kullanılarak eğitilmiş (*/trained*) görüşmeciler tarafından yürütülür. Ancak, anketin aksine, görüşme formu yanıtlayıcılara gösterilmeyen ve yalnızca görüşmecinin görebildiği özel talimatları ve görüşmecinin kişisel gözlemlerini ve yorumlarını kaydedebileceği bir alanı içerir. Buna ilaveten, posta yoklamalarının tersine, görüşmeci yanıtlayıcı tarafından ortaya koyulan meseleleri açıklamak ve irdelleyici sorular veya takip soruları sorma fırsatına sahiptir. Mamafih, görüşmeler zaman-alıcıdır ve kaynak-yoğundur. Görüşmecinin özel görüşme yeteneklerine sahip olması gerekir. Aynı zamanda, görüşmeci ölçüm enstrümanının bir parçası olarak addedilir ve gözlenen yanıtları yanlışlaştırmamak için proaktifçe çaba göstermelidir.

En tipik görüşme şekli kişisel veya **yüz-yüze görüşme**dir, öyle ki görüşmeci, sorular sormak ve verilen yanıtları kaydetmek için yanıtlayıcı ile doğrudan çalışır. Kişisel görüşmeler yanıtlayıcının evinde veya işyerinde yapılabilir. Bu yaklaşım hatta bazı yanıtlayıcılar tarafından tercih edilebilirken, diğer bazı yanıtlayıcılar özel alanlarına bir yabancı girmesinden rahatsızlık duyabilir. Mamafih, yetenekli görüşmeciler her durumda yanıtlayıcıları işbirliğine ikna edebilirler ve yanıt oranını dramatik olarak artırabilirler.

Kişisel görüşmenin bir tipi grup görüşmesidir ki **odak grup** olarak adlandırılır. Bu teknikte, aynı mekanda toplanan yanıtlayıcıların küçük bir grubu (genellikle 6-10 yanıtlayıcı) ile hep birlikte görüşme yapılır. Görüşmeci esasen bir ortam-yaratıcı rolündedir ki işi tartışmayı yönetmek ve herkesin bir yanıt verme fırsatını yakalamasını temin etmektir. Odak gruplar, yoklama araştırmasının başka şekillerine göre, kompleks meselelerin daha derin incelenmesine imkan sağlar çünkü insanlar başkalarının konuşmalarını duydukları zaman, daha önce düşünmedikleri yanıtlar veya fikirler için tetiklenmiş olurlar. Mamafih, odak grup tartışması dominant bir kişilik tarafından domine edilebilir ve bazı bireyler kendi akranlarının veya kendilerinden daha üstün bireylerin önünde, bilhassa mesela çalışanların performans-düşüklüğü veya ofis politikaları gibi duyarlı konular ele alınırken, görüşlerini seslendirmekten imtina edebilirler. Küçük örnek hacimlerinden dolayı, odak gruplar genellikle keşifçi araştırma için kullanılır, betimsel veya açıklayıcı araştırmadan daha ziyade.

Görüşme yoklamasının bir başka tipi **telefon görüşmeleridir**. Bu teknikte, görüşmeciler, standart yoklama sorularını sormak için, bir telefon rehberinden rastgele seçilen potansiyel yanıtlayıcılar ile telefonda konuşurlar. Son zamanlarda ortaya çıkan ve daha teknolojik bir yaklaşım kompüter-destekli telefon görüşmesidir (CATI), öyle ki görüşmeci bir telefon operatörüdür ve görüşme prosesi boyunca bir kompüter programı tarafından ekranda gösterilen talimatlar ve sorulacak sorular ile yönlendirilir. Akademide, hükümette ve ticarete yoklama araştırmacıları arasında bu tekniğin kullanımı artmaktadır. Bu sistem aynı zamanda bir rastgele numara arama tekniği kullanarak yanıtlayıcıları rastgele seçmekte ve bir ses yakalama tekniği kullanarak yanıtları kaydetmektedir. Yanıtlayıcıların telefona cevap vermeleri halinde, yüksek yanıt oranlarına ulaşılabilir. Bu teknik telefon yoğunluğunun düşük olduğu popülasyonlar için kullanılmamalıdır ve yine eğer iletişimde mesela grafikler veya ürün demonstrasyonları gibi sesten-başka enformasyon yer alacaksa, bu teknik kullanılmamalıdır.

Görüşmecinin rolü. Görüşmecinin görüşme prosesindeki rolü komplekstir ve çok yönlüdür, öyle ki şu görevleri ihtiva eder:

- **Görüşmeye hazırlanmak:** Veri toplama hareketinin ön cephesinde görüşmeci yer aldığı için, toplanan verinin kalitesi büyük oranda görüşmecinin vazifesini yapmak için ne kadar iyi eğitilmiş (*/trained*) olduğuna bağlıdır. Görüşmeci görüşme prosesi ve yoklama metodu hakkında eğitilmelidir (*/train*) ve aynı zamanda öğrenme-çalışmasının maksadına, yanıtların nasıl saklanacağına ve görüşmeci yanlılığına dair bilgisi

olmalıdır. Hakeza, formal öğrenme-çalışmasından evvel, görüşmeci prova yapmalı ve zamanını ayarlamalıdır.

- *Yanıtlayıcıların işbirliğini tesis etmek:* Bilhassa kişisel hane-içi yoklamalarda, görüşmeci spesifik adresleri tespit etmeli ve yanıtlayıcıların görüşmeye en istekli olabilecekleri yer ve zamanı tespit etmeye çaba göstermelidir. Bir görüşmeci aynı zamanda kendini öğrenme-çalışmasına katılma fikrini satan bir satıcı olarak görmelidir.
- *Yanıtlayıcıları motive etmek.* Yanıtlayıcılar çoğu zaman görüşmecinin motivasyonu ile beslenir. Eğer görüşmeci ilgisiz veya özensiz olursa, yanıtlayıcılar da yararlı veya enformasyonlu yanıtlar vermek için motive olmayacaklardır. Görüşmeci öğrenme-çalışması hakkında hevesini göstermeli, yanıtlayıcılara araştırmanın önemini anlatmalı ve görüşme boyunca yanıtlayıcıların ihtiyaçlarını gözetmelidir.
- *Herhangi bir kafa karışıklığını veya hassasiyeti açıklığa kavuşturmak.* Görüşmeciler yanıtlayıcılara empatiyle yaklaşmalı ve onların önceden akla gelmeyen hassasiyetlerini veya itirazlarını dikkate alarak, onları memnun etmeye çabalamalıdır. İlâveten, görüşmeci gerek duyduğu taktirde, görüşme formunda mevcut olmasa bile, irdeleyici sorular sormalıdır.
- *Yanıtın kalitesini temin etmeye çabalamak.* Görüşmeci toplanan verinin kalitesini yargılayabilecek en iyi pozisyonadadır ve alınan yanıtları, yanıtlayıcıların tavırlarına ve vücut dillerine ait kişisel gözlemlerini de kullanarak mütenasip bir şekilde destekleyebilir.

Görüşmeyi yürütmek. Görüşmeden önce, görüşmeci görüşme seansını icra etmek için bir paket hazırlamalıdır ki bu paketin içinde baş soruşturmacının veya sponsorun bir mektubu, yoklama enstrümanının yeterli sayıda kopyası, görüşmecinin fotoğraflı kimlik belgesi ve yanıtlayıcıların görüşmecinin gerçekliğini doğrulatabilecekleri bir telefon numarası yer almalıdır. Görüşmeci keza yanıtlayıcıları önceden arayarak eğer mümkünse bir randevu almaya çalışmalıdır. Görüşmeyi başlatırken, görüşmeci mesela “sizinle bir görüşme yapmama müsaade eder misiniz?” yerine “Çok önemli bir öğrenme-çalışması için görüşme yapmak üzere, bir kaç dakikanızı rica ediyorum” gibi imperatif ve güvenli bir ses tonuyla konuşmalıdır. Bir dakikadan daha az bir süre içerisinde, kendini tanıtmalı, göreviyle ilgili belgelerini takdim etmeli, öğrenme-çalışmasının maksadını 1-2 cümle ile açıklamalı ve yanıtlayıcıların yorumlarının ve gönüllü katılımlarının mahremiyetini garanti etmelidir. Büyük laflar veya jargon ile konuşmamalı ve spesifik olarak istenilmediği taktirde detaylara girmemelidir. Eğer görüşmeci görüşmeyi sesli olarak kaydetmek isterse, bunu yapmadan önce yanıtlayıcının sarahaten iznini istemelidir. Görüşme kaydediliyor olsa dahi, görüşmeci kilit meseleleri, irdelemeleri veya tabirleri not etmelidir.

Görüşme sırasında, görüşmeci soru formunu takip etmeli ve soruları tam olarak yazılmış olduğu üzere sormalı ve soruyu daha sempatik hale getirmek için kelimeleri kendince değiştirmemelidir. Hakeza, soruların sırasını değiştirmemeli veya hangi düşünceyle olursa olsun herhangi bir soruyu atlamamalıdır. Sorulara dair meseleler fiili görüşme seansından önce yapılacak prova esnasında tartışılmalıdır. Görüşmeci yanıtlayıcının cümlelerini bitirmemelidir. Eğer yanıtlayıcı çok kısa ve üstünkörü bir cevap verirse, görüşmeci daha düşünceli, eksiksiz bir yanıt vermesi için yanıtlayıcıyı irdelemelidir. Bazı irdeleme teknikleri şunlardır:

- *Sessiz irdeleme:* Yalnızca durmak ve beklemek (bir sonraki soruya geçmeksizin) yanıtlayıcıya daha detaylı bir yanıt vermesinin beklendiği mesajını verebilir.
- *Cesaretlendirme:* Ara sıra “ya” veya “öyle mi?” gibi karşılıklar vermek yanıtlayıcıyı daha fazla detaya girmesi için cesaretlendirebilir. Mamafih, görüşmeci yanıtlayıcının söylediği herhangi bir şeyi onayladığını veya onaylamadığını gösteren bir ifadede bulunmamalıdır.
- *De taylandırma isteme:* Mesela “Bunu daha detaylandırabilir misiniz?” veya “Az önce,

lisedeki bir deneyiminizden söz ettiniz. Onunla alakalı daha fazla konuşabilir misiniz?” gibi.

- **Yansıtma:** Görüşmeci, yanıtlayıcının söylediği şeyi tekrar ederek, psiko-terapistlerin kullandığı numarayı kullanmayı deneyebilir. Mesela, “Yaşadığınız deneyimi çok travmatik bulduğunuzu söylüyorsunuz.” dedikten sonra, durup, yanıtlayıcının ayrıntılara girmesini bekleyebilir.

Yoklama Araştırmasında Yanlılıklar

Tüm güçlü ve avantajlı yönlerine rajmen, yoklama araştırması çoğu zaman sistematik yanlılıklar ile mâlûldür ki bu yanlılıklar böylesi yoklamalardan türetilen bazı çıkarımları geçersizleştirebilir. Böyle beş yanlılık yanıt-yok yanlılığı, örnekleme yanlılığı, sosyal beğenilirlik yanlılığı, anımsama yanlılığı ve müşterek metot yanlılığıdır.

Yanıt-yok yanlılığı. Yoklama araştırması genellikle düşük yanıt oranları ile anılır. Bir posta yoklamasında, hatta iki veya üç anımsatmadan sonra dahi, tipik yanıt oranı %15-20 dir. Eğer hedeflenen yanıtlayıcıların büyük çoğunluğunun yoklamayı yanıtlamasında başarısız olunursa, o zaman yanıt-vermeyenlerin yanıt-vermemelerinin öğrenme-çalışmasının sonuçlarının geçerliliğinin sorgulanmasına neden olacak sistematik bir sebebi olup olmadığı haklı olarak düşünülmelidir. Mesela, memnuniyetsiz müşteriler, memnun müşterilere göre, yaşadıkları deneyim hakkında konuşmaya daha fazla isteklidir ve bu yüzden anket yoklamalarını veya görüşme yoklamalarını yanıtlamaya daha fazla yatkınlardır. Bundan ötürü, yanıtlayıcıların herhangi bir örneğinde memnuniyetsiz müşterilerin orantısı, içinden çekilmiş olduğu altta-yatan popülasyona göre daha yüksek olabilecektir. Bu halde, yalnızca sonuçlar genelleştirilebilirliklerini kaybetmekle kalmayacak aynı zamanda gözlenen çıkımlar yanlı örneğin yapıntısı olabilecektir. Yanıt oranlarını iyileştirmek için istihdam edilebilecek birkaç strateji şunlardır:

- **Önceden haber vermek:** Yürütülecek bir yoklamaya katılımlarının rica edildiği bir mektubu hedeflenen yanıtlayıcılara önceden göndermek onları önceden yoklamaya hazırlayabilir ve yanıt verme arzularını artırabilir. Mektupta öğrenme-çalışmasının maksadı ve önemi, veri toplama şekli (mesela, telefon görüşmesi ile, postalanacak bir yoklama formu ile) ve yapacakları işbirliği için duyulacak minnettarlık ifade edilmelidir. Bu tekniğin başka bir çeşidinde, yanıtlayıcıdan, öğrenme-çalışmasına katılıp katılamayacağını, temin edilen bedeli-önceden-ödenmiş bir posta kartı vasıtasıyla, bildirmesi rica edilebilir.
- **İçeriğin münasebeti:** Eğer bir yoklama yanıtlayıcılar için münasebeti veya önemi olan meseleleri inceliyor ise, böyle bir yoklamayı yanıtlamaya daha fazla meyilli olabilirler.
- **Yanıtlayıcı-dostu anket:** Daha kısa yoklama anketleri, daha uzun olanlarına göre daha yüksek yanıt oranlarına ulaşma eğilimi gösterir. Dahası, açık, mütecaviz-olmayan ve yanıtlaması kolay sorular daha yüksek yanıt oranları temin etmek eğilimindedir.
- **Onaylatma:** Organizasyonel yoklamalarda, öğrenme-çalışmasının organizasyon için önemini olurlayan bir üst yöneticinin onayını almak yararlıdır. Böylesi bir onay bir kapak veya sunum mektubu formunda olabilir ve yanıtlayıcıların gözünde araştırmacının kredibilitesini artırır.
- **Talebi yenilemek:** Yanıt-vermeyenlerden yoklamaya katılmalarını birden çok kere yeniden talep etmek yanıt-vermeyenlerin bazılarını yanıt vermeye ikna edebilir, yanıtlarını geç verseler dahi.
- **Görüşme eğitimi:** Görüşme talebinde nasıl bulunacağı, potensiyel yanıtlayıcıları belirlemek için kompüterize telefonla-arama tekniklerini nasıl kullanacağı ve ulaşılamayan yanıtlayıcılar için geri dönüşleri nasıl planlayacağı konusunda eğitilmiş ve becerikli görüşmeciler ile görüşmelerin yanıt oranları artırılabilir.

- *Teşvikler*: Nakit para, hediye kartı veya mesela kalem, stres topu, piyango bileti, indirim kuponu veya katılımcı adına bir hayır kurumuna bağış gibi eşantyonlarla, en azından bazı popülasyonlarda yanıt oranları artırılabilir.
- *Parasal-olmayan teşvikler*: Bilhassa firmalar finansal teşviklerden ziyade parasal-olmayan teşviklere daha meyillidir. Mesela firmanın kendi yanıtını yoklamadaki tüm yanıtlar ile karşılaştıran bir *benchmarking* raporu bunun bir misalidir.
- *Mahremiyet ve gizlilik*: Son olarak, yanıtlayıcıların özel verisinin veya yanıtlarının başkalarının eline geçmeyeceğine dair verilecek güvence yanıt oranlarının artmasına yardımcı olabilir.

Örnekleme yanlılığı. Kamuya açık telefon numaralarının bir rastgele örneğini aramak suretiyle yürütülen telefon yoklamaları rehberde yer almayan telefon numaralarını, mobil telefonları ve yoklama esnasında telefona cevap veremeyenleri sistematik olarak dışlayacaktır ve rehberde numarası yer alan kablolu telefon kullanıcılarını ve günün çoğunu telefonun bağlı olduğu mekanda geçirenleri (mesela işsizler, engelliler veya yaşlılar) orantısız bir sayıda ihtiva edecektir. Hakeza, çevrimiçi yoklamalar devamlı internette olan öğrencileri ve gençleri yine orantısız bir sayıda ihtiva etmeye meyilliyken, diğer yandan mesela fakirler ve yaşlılar gibi kompütürlere ve internete ulaşımı kısıtlı olan insanları sistematik olarak dışlamaya meyillidir. Benzerce, anket yoklamalarının da soruları okuyamayacak, anlayamayacak veya manalı yanıtlar verebilmeyecek çocukları ve okur-yazar olmayanları dışlama eğilimi vardır. Örnekleme yanlılığının farklı bir tipi de, mesela öğretmenlere (veya ebeveynlere) öğrencilerinin (veya çocuklarının) akademik başarılarını sormak veya CEO'lara şirketlerinin operasyonel ayrıntılarını sormak gibi yanlış popülasyonu örneklemekle ilgilidir. Böyle yanlılıklar yanıtlayıcıların örneğini niyet-edilen popülasyonun temsilcisi olmaktan uzaklaştırır ve yanlı örnekten yapılacak çıkarımların genelleştirilebilirliğini zedeler.

Sosyal beğenilirlik yanlılığı. Yanıtlayıcıların ekseriyeti kendileri, işverenleri, aileleri veya arkadaşları hakkında negatif görüşlerden veya utandırıcı yorumlardan kaçınmaya meyillidir. Mesela, “Proje takımınızın disfonksiyonel olduğunu düşünüyor musunuz?”, “İşyerinizde ofis-içi siyaset yoğun mudur?” veya “İnternette müzik dosyalarını yasadışı olarak hiç indirdiniz mi?” gibi negatif sorulara, araştırmacı doğru-sözlü yanıtlar alamayabilir. Kendilerini sosyal olarak beğenilir kılmak adına, yanıtlayıcıların bu “gerçeği ters-yüz etme” eğilimi “sosyal beğenilirlik yanlılığı” olarak adlandırılır ki bu yanlılık yoklama araştırmasından elde edilen yanıtların geçerliliğini zedeler. Bir anket yoklamasında, sosyal beğenilirlik yanlılığının üstesinden gelmenin pratikte hiç bir yolu yoktur. Fakat bir görüşme ortamında ferasetli bir görüşmeci tutarsız cevapları tespit edebilir ve yanıtlayıcının yorumlarını desteklemek üzere irdeleyici sorular sorabilir veya kişisel gözlemlerini kullanabilir.

Anımsama yanlılığı. Yoklama sorularına alınacak yanıtlar çoğu zaman deneklerin belleklerine ve yanıtlama yeteneklerine bağlıdır. Bilhassa uzak geçmişte cerayan etmiş olaylar söz konusu olduğu zaman, yanıtlayıcılar o zamanki motivasyonlarını veya davranışlarını kafi derecede anımsamayabilirler veya geçen zaman zarfında belki de o olaylara ait bellekleri evrilmiş olabilir ve artık ulaşılabilir olmayabilir. Mesela, eğer yanıtlayıcılardan bir yıl önce kullandıkları kompüter teknolojisi veya hatta doğum günleri gibi çocukluk anılarını betimlemeleri istenirse, yanıtları isabetli olmayabilir, anımsama zorluklarından ötürü. Anımsama yanlılığının üstesinden gelmenin olanaklı bir yolu onlardan belleklerindeki algıları veya motivasyonları anımsamalarını istemek yerine, yanıtlayıcıların belleklerini o olayların cereyan ettiği zamanki spesifik olaylara çıpalamaktır.

Müşterek metot yanlılığı. Müşterek metot yanlılığı ile zamanın aynı noktasında ölçülen açıklayıcı değişken ve yanıt değişkeni arasında paylaşılan suni kovaryansın miktarı kastedilir. Bir kesitsel (/cross-sectional) yoklamada hem açıklayıcı değişkeni hem de yanıt değişkenini

ölçmek için mesela bir anket gibi aynı enstrümanı kullanmak buna bir misaldir. Böyle durumlarda, soruşturulan fenomen ölçüm yapıntılarından kifayyetle ayrılamayabilir. Müşterek metot yanlışlığını test etmek için standart istatistiksel testler mevcuttur ki bunların arasında Harmon'un tek-faktör testini (Padsakoff ve ark. 2003) ve Lindell ve Whitney'in (2001) market değişken tekniğini sayabiliriz. Bu tip yanlışlıklardan kaçınmak için açıklayıcı değişken ve yanıt değişkeni boylamsal bir yoklama dizaynı kullanılarak zamanın farklı noktalarında ölçülebilirler veya bu değişkenleri ölçmek için farklı metotlar kullanılabilir.

Bölüm 10

Deneyisel Araştırma

Deneyisel araştırma tüm araştırma dizaynları içerisinde en sıkı olanıdır ve çoğu zaman araştırma dizaynlarının “altın standardı” olarak mülahaza edilir. Bu dizaynda, bir veya birden çok sayıdaki açıklayıcı değişken araştırmacı tarafından manipüle edilir (muameleler olarak), denekler (*/subjects*) açıklayıcı değişkenin farklı düzeylerine (yani muamelelere) rastgele atanır ve muamelelerin bir veya birden çok yanıt değişkenine ait çıkımlar üzerindeki sonuçları gözlenir. Deneyisel araştırmanın kendine has gücü onun **içsel geçerliliğidir** (nedensellik/*causality*). Bir deney dizaynının bunu yapabilmesini sağlayan şey, harici değişkenlerin suni etkilerini kontrol altına almak suretiyle, muamele manipülasyonu vasıtasıyla neden ve etki arasında bir bağ kurabilme kabiliyetidir.

Deneyisel araştırma, öğrenme-çalışmasının amacının neden-ve-etki ilişkilerini incelemek olan açıklayıcı araştırma için (betimsel veya keşifsel araştırmadan daha ziyade) en uygundur. Manipüle edilebilecek veya kontrol altına alınabilecek nispeten sınırlı sayıda açıklayıcı değişken ihtiva eden araştırmalar için de yine uygundur. Deneyisel araştırma laboratuarda veya sahada yürütülebilir. Laboratuarda kurulan (yapay) ve yürütülen **laboratuvar deneylerinin** genellikle *içsel geçerliliği* yüksek olmakla beraber, *dışsal geçerliliği* düşüktür çünkü öğrenme-çalışmasının içinde yürütüldüğü yapay kurulum (laboratuvar) gerçek dünyayı yansıtmayabilir. Mesela gerçek bir organizasyon gibi bir saha ortamında yürütülen **saha deneylerinin** hem içsel hem de dışsal geçerliliği yüksektir. Ancak, böyle deneyler nispeten nadirdir çünkü bir saha kurulumunda muameleleri manipüle etmek ve harici etkileri kontrol altına almak zordur.

Deneyisel araştırma iki geniş kategoride gruplandırılabilir: hakiki deneyisel dizaynlar ve yarı-deneyisel (*/quasi-experimental*) dizaynlar. Her iki grupta yer alan dizaynlar muamele manipülasyonunu gerektirmekle beraber, yalnızca deneyler rastgele atamayı gerektirir. Yarı-deneyler rastgele atamayı gerektirmez.

Bazen, deneyisel-olmayan araştırma (*/non-experimental*) tabiri kullanılır ki bununla gerçekte bir deney dizaynı kastedilmez fakat muamele manipülasyonunu veya rastgele atamayı ihtiva etmeyen mesela yoklama araştırması, gözlemsel araştırma ve korelasyonel öğrenme-çalışmaları gibi başka her türden araştırmayı kapsayıcı bir tabir olarak kullanılır.

Temel Kavramlar

Muamele ve kontrol grupları. Deneyisel araştırmada, deneklerin bir kısmı **muamele** adı verilen bir (veya birden çok sayıda) deneyisel stimülasyona maruz bırakılırken (**muamele grubu**), diğer denekler böyle bir stimülasyona maruz bırakılmaz (**kontrol grubu**). Eğer muamele grubundaki deneklerin, kontrol grubundaki deneklere göre, çıkım değişkenleri bakımından daha farklı oldukları gözlenirse, muamele başarılı addedilebilir. Deneyisel stimülusun çoklu düzeyleri de uygulanabilir. Bu halde, birden çok sayıda muamele grubu oluşur. Mesela demans gibi belli bir tıbbi durumu tedavi etmek maksadıyla yeni bir ilacı test etmek için, demans hastalarının bir örneği eğer rastgele üç gruba bölünerek, birinci gruba ilacın yüksek bir dozu, ikinci gruba ilacın düşük bir dozu ve üçüncü gruba mesela şeker hapı gibi bir plasebo verilir ise, ilk iki grup deney grupları ve üçüncü grup kontrol grubudur. Bir zaman periyodu boyunca ilaç verildikten (yani, muameleler uygulandıktan) sonra, eğer deney grubundaki denekler kontrol grubundaki deneklere göre anlamlıca daha fazla iyileşirler ise, ilacın etkili olduğunu söyleyebiliriz. Aynı zamanda, yüksek ve düşük dozlu deney gruplarının

durumlarını karşılaştırmak suretiyle, yüksek dozun düşük dozdan daha etkili olup olmadığını belirleyebiliriz.

Muamele manipülasyonu. Muameleler deneysel araştırmanın yekta hususiyetidir ki bir deneysel araştırmayı başka tüm araştırmalardan ayrı kılar. Muamele manipülasyonu neden-ve-etki ilişkilerinde “nedeni” kontrol altına almayı sağlar. Doğal olarak, deneysel araştırmanın geçerliliği muamelenin nasıl manipüle edildiğine bağlıdır. Muamele manipülasyonu, deneysel öğrenme-çalışmasından önce, öntestler ve pilot testler kullanılarak denetlenmelidir. Muamelenin uygulanmasından önce yapılan ölçümlere **öntest ölçüleri** adı verilirken, muamelenin uygulanmasından sonra yapılan ölçümlere **sontest ölçüleri** adı verilir.

Rastgele seçim ve rastgele atama. Rastgele seçim bir popülasyondan veya örnekleme çerçevesinden bir örneği rastgele çekme prosedürüdür. Bu yaklaşım tipik olarak yoklama araştırmasında istihdam edilir ve popülasyondaki her bir ünitenin örneğe seçilmek için pozitif bir şansa sahip olmasını emniyet altına alır. Rastgele atama, mamafih, deney ünitelerini deney gruplarına veya kontrol grubuna rastgele atama prosedürüdür. Bu, hakiki deneysel araştırmada standart bir uygulamadır ve muamelenin uygulanmasından önce muamele gruplarının birbirine ve kontrol grubuna benzer (eşit) olmasını temin eder. Rastgele seçim, örnekleme ile alakalıdır ve bu bakımdan, bulguların dışsal geçerliliği (genelleştirilebilirlik) ile yakından alakalıdır. Mamafih, rastgele atama dizayn ile alakalıdır ve bu bakımdan, içsel geçerlilik ile en çok alakalıdır. Hem rastgele seçime hem de rastgele atamaya beraber sahip olmak iyi dizayn edilmiş bir deneysel araştırmada mümkün olur. Yarı-deneysel araştırma ne rastgele seçimi ne de rastgele atamayı ihtiva eder.

İçsel geçerliliğe yönelik tehditler. Her ne kadar deneysel dizaynların, çıkarımlarının içsel geçerliliği bakımından, başka araştırma metodlarına göre daha sıkı oldukları mülahaza ediliyor olsa bile (muamele manipülasyonu marifetiyle nedenleri kontrol altına alma kabiliyetine sahip olmaları hasebiyle), onlar da içsel geçerliliklerine yönelik tehditlere karşı bağışık değildir. İçsel geçerliliğe yönelik tehditlerden bazıları aşağıda lise öğrencilerinin matematik yeteneklerini ıslah etmek üzere uygulanan özel bir iyileştirici matematik kurs programının etkisini ölçen bir öğrenme-çalışması bağlamı içerisinde betimlenmektedir.

- *Tarih tehdidi* gözlenen etkilere (yanıt değişkenleri) deneysel muameleden daha ziyade harici veya tarihsel olayların neden olması olanağıdır. Mesela, öğrencilerin iyileştirici kurs programı sonrası skorlarındaki yükselmeye ayrıca kendi okullarında matematik sınavına hazırlanmaları neden olmuş olabilir, iyileştirici kurs programının kendisinden daha ziyade.
- *Olgunlaşma tehdidi* şu olanaklılığa tekabül eder ki gözlenen etkilere deneklerin doğal olgunlaşmaları (yani kompleks kavramları anlamada entellektüel yeteneklerindeki genel bir iyileşme) neden olur, deneysel muamelenin kendisinden daha ziyade.
- *Testetme tehdidi* öntest-sontest dizaynlarına mahsus bir tehdittir öyle ki deneklerin sontest yanıtları öntest yanıtlarına koşullandırılmıştır. Mesela, eğer öğrenciler öntestte verdikleri cevapları anımsıyorlar ise o cevapları sontestte de tekrarlama eğilimi gösterebilirler.
- *Enstrümantasyon tehdidi* yine öntest-sontest dizaynlarına mahsustur ve şu olanaklılığa izafe edilir ki öntest puanları ve sontest skorları arasındaki fark iyileştirici matematik kursu programından kaynaklanmamaktadır, fakat uygulanan testlerdeki mesela daha yüksek veya daha düşük zorluk derecelereine sahip olmaları gibi değişikliklerden kaynaklanmaktadır.
- *Mortalite tehdidi* şu olanaklılığa matuftur ki denekler öğrenme-çalışmasına ait muamele ve kontrol gruplarından sistematik bir sebepten dolayı diferansiyel oranlarda eksilebilirler, mesela eksilen bu öğrencilerin çoğunlukla öntestten düşük skor yapan

olması gibi. Eğer performansı düşük öğrenciler eksiltirirse, son testin sonuçları geriye kalan yüksek performanslı öğrencilerin ekseriyeti tarafından yapay olarak şişirilmiştir olacaktır.

- *Regresyon tehdidi*, aynı zamanda aritmetik-ortalamaya gerileme (/regression to the mean) de denir, bir grubun son testte ölçülen genel performansının, öngörüldüğü yönden daha ziyade, o ölçümlerin aritmetik-ortalamasına gerilemesiyle görülen istatistiksel eğilime tekabül eder. Mesela, eğer denekler öntestte yüksek skorlar yapmışlar ise, son testte daha düşük skorlar (aritmetik-ortalamaya daha yakın) yapma eğilimi göstereceklerdir, çünkü bunların öntestte yaptıkları yüksek skorlar (aritmetik-ortalamadan uzak) mümkündür ki bir istatistiksel aberasyondur. Bu problem rastgele olmayan örneklerde ve iki ölçümün tam korelasyonlu olmadığı durumlarda daha yaygınca görülme eğilimindedir.

İki-Gruplu Deney Dizaynları

En basit hakiki deney dizaynları iki gruplu dizaynlardır ki bir muamele grubu ve bir kontrol grubu ihtiva ederler. Bir muamele olarak manipüle edilebilecek tek bir açıklayıcı değişkenin etkisini test etmek için bunlar en uygun dizaynlardır. İki-gruplu dizaynların başlıca iki tipi öntest-son test kontrol grubu dizaynı ve salt-son test kontrol grubu dizaynıdır, buna ilaveten kovaryans dizaynlarını ihtiva eden çeşitleri de mevcuttur. Bu dizaynların gösteriminde çoğu zaman standartlaştırılmış bir dizayn notasyonu kullanılır; şöyle ki R deneklerin gruplara rastgele atanmasını temsil eder, X muamele grubuna uygulanan muameleyi temsil eder ve O yanıt değişkeninin öntest ve son test gözlemlerini temsil eder (muamele ve kontrol gruplarının öntest ve son test gözlemleri arasında ayırım yapmak için farklı alt-indices kullanılarak).

Öntest-son test kontrol grubu dizaynı. Bu dizaynda, denekler muamele ve kontrol gruplarına rastgele atanır, denekler üzerinde ilgilenilen yanıt değişkeninin başlangıç (öntest) ölçümleri yapılır, muamele grubuna bir muamele (ilgilenilen açıklayıcı değişkeni temsil eden) uygulanır ve yanıt değişkeni yeniden ölçülür (son test). Bu dizaynın notasyonu Figür 10.1 de gösterilmektedir.

R	O ₁	X	O ₂	(Muamele grubu)
R	O ₃		O ₄	(Kontrol grubu)

Figür 10.1. Öntest-son test kontrol grubu dizaynı

Öntest-son test dizaynında deneysel muamelenin etkisi E muamele ve kontrol grupları arasında son test ve öntest skorlarının farkı olarak ölçülür:

$$E = (O_2 - O_1) - (O_4 - O_3)$$

Bu dizaynın istatistiksel analizi muamele ve kontrol grupları arasında t-testi veya varyans analizi (ANOVA) ile yapılabilir. Öntest-son test dizaynı içsel geçerliliğe yönelik çeşitli tehditleri elemine edebilir, mesela maturasyon, testetme ve regresyon gibi. Çünkü bu tehditlerin hem muamele grubuna hem de kontrol grubuna genel olarak benzer (rastgele) tarzda tesir edeceği beklenebilir. Seçim tehdidi rastgele atama marifetiyle kontrol altına alınır. Mamafih, içsel geçerliliğe yönelik başka tehditler yine mevcut olabilir. Mesela, eğer iki gruptaki eksilme oranları farklıysa, mortalite bir problem olabilir ve öntest ölçümü son test ölçümünü yanıltabilir (bilhassa eğer öntest sıradışı mevzuları veya içeriği haizse).

Salt-son test kontrol grubu dizaynı. Bu dizayn öntest-son test dizaynının daha basit bir çeşididir öyle ki öntest ölçümleri dahil edilmez. Dizayn notasyonu Figür 10.2 de gösterilmektedir.

R	X	O ₁	(Muamele grubu)
R		O ₂	(Kontrol grubu)

Figür 10.2. Salt-sontest kontrol grubu dizaynı

Deneyisel muamelenin etkisi E basitçe muamele ve kontrol grupları arasında sontest skorlarının farkı olarak ölçülür:

$$E = O_2 - O_1$$

Bu dizayn için mütenasip istatistiksel analiz metodu yine muamele ve kontrol grupları arasında t -testi veya varyans analizi (ANOVA) olabilir. Bu dizaynın basitliği, içsel geçerlilik bakımından kendini öntest-sontest dizaynına göre daha çekici kılar. Bu dizayn maturasyon, testetme, regresyon, seçim ve öntest-sontest interaksyonu tehditlerini kontrol altına alır fakat mortalite tehdidi mevcudiyetini devam ettirebilir.

Kovaryans dizaynları. Bazen, yanıt değişkenine ait ölçümler *kovaryete* adı verilen harici değişkenlerin tesirinde kalabilir. Kovaryeteler öyle değişkenlerdir ki bir deneysel öğrenme-çalışmasında merkezi bir ilgiye mazhar değillerdir, fakat bir araştırma dizaynında, yanıt değişkeni üzerindeki potansiyel etkilerini elemine etmek için, kontrol altına alınmaları lazım gelen değişkenlerdir. Böylece, ilgilenilen açıklayıcı değişkenin yanıt değişkeni üzerindeki etkisi daha isabetlice belirlenebilecektir. Öncesinde tartıştığımız deney dizaynları, böylesi kovaryeteleri kontrol altına almazlar. Kovaryans dizaynı (aynı zamanda konkomitant değişken dizaynı olarak ta adlandırılır) öntest-sontest kontrol grubu dizaynının bir özel tipidir şöyle ki öntest ölçüsü, yanıt değişkeninin bir ölçümü değil de, ilgilenilen kovaryetenin bir ölçümüdür. Kovaryans dizaynı notasyonunu Figür 10.3 te görebiliyoruz ki orada C kovaryeteyi temsil etmektedir:

R	C	X	O ₁	(Muamele grubu)
R	C		O ₂	(Kontrol grubu)

Figür 10.3. Kovaryans dizaynı

Muamelenin etkisi E muamele ve kontrol grupları arasında kovaryeteye göre düzeltilmiş sontest skorlarının farkı olarak ölçülür:

$$E = O_2(C) - O_1(C)$$

Kovaryetenin mevcudiyetinden dolayı, bu dizayn için mütenasip istatistiksel analiz metodu kovaryans analizi (ANCOVA) olmalıdır. Bu dizayn salt-sontest dizaynının tüm avantajlarına sahip olmakla beraber, aynı zamanda kovaryeteleri kontrol altına alması hasebiyle içsel geçerliliği de sahiptir.

Faktöriyel Dizaynlar

İki-gruplu dizaynlar kifayetsizdir eğer araştırmanız iki veya daha çok sayıda açıklayıcı değişkenin manipülasyonunu gerektiriyor ise. Bu durumda, dört veya daha çok sayıda muamele grubu oluşacaktır. Böylesi dizaynlar faktöriyel deneysel araştırmada hayli popülerdir ve müştereken faktöriyel dizaynlar olarak adlandırılırlar. Bu dizaynlarda her bir açıklayıcı değişkene bir *faktör* denir ve bir faktörün her bir alt-bölümüne bir *düzey* denir. Böylece faktörlerin düzeyleri ve bunların kombinasyonları birer muamele grubu oluşturur. Faktöriyel dizaynlar araştırmacının yanıt değişkeni üzerine her bir açıklayıcı değişkenin bireysel etkisini (ana etkiler olarak adlandırılır) incelemesini temin etmesinin yanı sıra, aynı zamanda bu

açıklayıcı değişkenlerin ortak etkisini (interaksiyon etkisi olarak adlandırılır) de incelemesini temin eder.

En temel faktöriyel dizayn 2x2 faktöriyel dizayndır ki iki muameleden ibarettir, öyle ki bu muamelelerin her birinin iki düzeyi vardır (mesela, yüksek/düşük veya var/yok gibi). Mesela, diyelim ki iki farklı tipte öğretim tekniğinin (sınıf-içi ve çevrim-içi öğretim) öğrenme çıktımlarını karşılaştırmak istiyoruz ve aynı zamanda bu etkilerin öğretim süresi (haftada 1.5 veya 3 saat) ile değişip değişmediğini de incelemek istiyoruz. Bu durumda, iki faktörümüz var: öğretim tipi ve öğretim süresi; her birinin iki düzeyi var (öğretim tipi için sınıf-içi ve çevrim-içi ve öğretim süresi için haftada 1.5 saat ve 3 saat), ki Figür 10.4 de gösterilmektedir. Eğer öğretim süresine üçüncü bir düzeyi de eklemeyi arzu edersek (diyelim haftada 6 saat), o zaman ikinci faktörümüz üç düzeyli olur ve 2x3 faktöriyel bir dizayna sahip oluruz. Öte taraftan, eğer mesela grup çalışması (var/yok) gibi üçüncü bir faktörü daha eklemeyi arzu edersek, 2x2x2 faktöriyel bir dizayna sahip oluruz. Bu notasyonda, her bir sayı bir faktörü temsil etmektedir ve her bir faktörün değeri o faktörün düzeylerinin sayısını temsil etmektedir.



Figür 10.4. 2x2 faktöriyel dizayn

Faktöriyel dizaynlar aynı zamanda bir dizayn notasyonu vasıtasıyla da gösterilebilirler, Figür 10.4 ün sağ panelinde görüldüğü gibi. R deneklerin muamele gruplarına rastgele atanmasını temsil ediyor, X muamele gruplarının kendilerini temsil ediyor (X in alt indisi faktörlerin düzeylerini temsil ediyor) ve O yanıt değişkenine ait gözlemleri temsil ediyor. Dikkatini çekmek istiyorum ki 2x2 faktöriyel dizayn dört muamele grubuna sahiptir ki bunlar iki faktörün ikişer düzeyinin dört kombinasyonudur. Aynı minvalde, 2x3 faktöriyel bir dizayn altı muamele grubuna sahip olacaktır ve 2x2x2 faktöriyel bir dizayn sekiz muamele grubuna sahip olacaktır. Bir kural olarak söylenebilir ki bir faktöriyel dizaynın her bir hücrelerinin örnek hacmi minimum 20 olmalıdır (bu kestirim Cohen'in orta etki miktarına dayalı güç hesaplarından türetilir). Böylece, 2x2x2 faktöriyel bir dizayn toplamda 160 deneklik minimum bir örnek hacmi gerektirir ki her bir hücrede en az 20 denek demektir. Görüldüğü üzere, bir faktöriyel dizaynda faktörlerin sayısı veya faktör düzeylerinin sayısı artarsa, veri toplama maliyeti de esaslı olarak artabilmektedir. Bazen, kaynak kısıtlarından dolayı, böylesi faktöriyel dizaynlarının bazı hücrelerindeki muameleler uygulanmayabilir ki bunlara *eksik faktöriyel dizaynlar* adı verilir. Böyle eksik dizaynlar araştırmacının eksik faktörler hakkında netice çıkarma kabiliyetini zaafa uğratar.

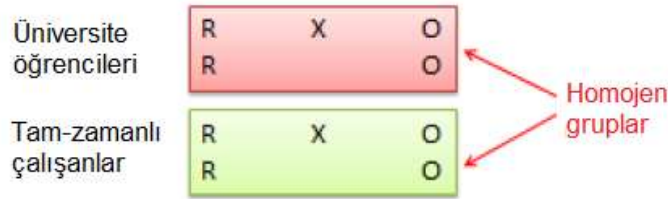
Bir faktöriyel dizaynda, eğer yanıt değişkeni, diğer faktörlerin tüm düzeylerinde, bir faktörün düzeyleri arasında anlamlı bir fark gösteriyorsa, bir **ana etkinin** mevcut olduğu söylenir. Yanıt değişkeninin faktör düzeyleri arasında hiç bir fark göstermemesi bazçizgisi olarak alınarak bir ana etki değerlendirilir. Yukarıdaki misalde, öğretim tipinin veya öğretim zamanının öğrenim çıktımları üzerine bir ana etkisini görebilirdik; her ikisi de bir ana etkiye sahip olabilirdi. Bir

faktörün düzeyleri arasında, yanıt değişkeninin gösterdiği farklılık eğer bir diğer faktörün düzeylerine göre değişiyorsa, bir **interaksiyon etkisinin** mevcut olduğu söylenir. Bizim misalimizde, eğer öğrenme çıkımları üzerine öğretim tipinin etkisi haftada 3 saatlik öğretim süresi için, haftada 1.5 saatlik öğretim süresine göre daha büyükse, o zaman diyebiliriz ki öğrenim çıktıları üzerine öğretim tipi ve öğretim süresi arasında bir interaksiyon etkisi mevcuttur. Dikkat etmelisin ki interaksiyon etkisinin mevcudiyeti ana etkileri baskılar ve onları münasebetsiz kılar. Eğer interaksiyon etkisi anlamlı bulunursa, ana etkilerin interpretasyonunu yapmanın bir manası kalmaz.

Melez Deney Dizaynları

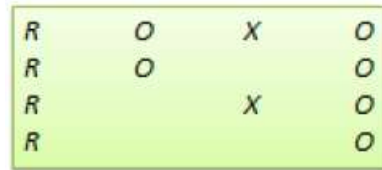
Melez dizaynlar öyle dizaynlardır ki daha müesses dizaynların vasıfları biraraya getirilmek suretiyle teşkil edilirler. Böylesi üç melez dizayn rastgeleleştirilmiş blok dizaynı, Solomon dört-grup dizaynı ve çapraz tekraralama dizaynı.

Rastgeleleştirilmiş blok dizaynı. Bu, salt-sontest veya öntest-sontest kontrol grubu dizaynlarının bir çeşididir ki bu tip dizaynlarda denek popülasyonu *blok* adı verilen nispeten homojen altgruplara ayrılabilir ki bu altgrupların her birinde deney tekrarlanır. Mesela, aynı salt-sontest dizaynını üniversite öğrencileri ve tam-zamanlı çalışan meslek sahipleri (kendi içlerinde homojen iki blok) arasında tekrarlamak istersen, her iki bloktaki denekler muamele grubu ve kontrol grubu arasında rastgele bölünür (Figür 10.5 e bak). Bu dizaynın maksadı bloklar arasındaki farklara hamledilebilecek “paraziti” veya verideki varyansı azaltmaktır öyle ki ilgilenilen fiili etki daha isabetlice saptanabilsin.



Figür 10.5. Rastgeleleştirilmiş blok dizaynı

Solomon dört-grup dizaynı. Bu dizaynda, örnek iki muamele grubuna ve iki kontrol grubuna bölünür. Bir muamele grubu ve bir kontrol grubu önteste tabi tutulur ve diğer iki grup önteste tabi tutulmaz. Bu dizayn salt-sontest dizaynı ve öntest-sontest kontrol grubu dizaynının bir kombinasyonunu temsil eder ve öntest ölçümünün sontest ölçüleri üzerindeki potansiyel yanıltıcı etkisini test etmek için kullanılır. Söz konusu etki öntest-sontest dizaynlarında ortaya çıkma eğilimi gösterir fakat salt-sontest dizaynlarında değil. Dizayn notasyonu Figür 10.6 da gösterilmektedir.



Figür 10.6. Solomon dört-grup dizaynı

Çapraz tekraralama dizaynı. Bu, iki safhada, üç ölçüm dalgası ile tatbik edilen bir iki-grup

dizaynıdır. Figür 10.7 de gösterildiği üzere, birinci safhadaki muamele grubu ikinci safhada kontrol grubu olarak hizmet eder ve birinci safhadaki kontrol grubu ikinci safhada muamele grubu olur. Başka bir deyişle, orijinal dizayn gruplar arasında çaprazlanan muamele/kontrol rolleriyle zamanda yinelenir veya tekrarlanır. Bu öğrenme-çalışmasının sonucunda, tüm katılımcılar ya birinc safhada ya ikinci safhada muameleye tabi tutulmuş olacaktır. Bu dizayn, organizasyonel programların (mesela, çalışan eğitimi) safhalar halinde tatbik edildiği veya düzenli aralıklarla yinelandığı organizasyonel bağlamlarda en fizibıldır.

R	O	X	O		O
R	O		O	X	O

Figür 10.7. Çapraz tekrarlama dizaynı

Yarı-deneysel Dizaynlar

Yarı-deneysel dizaynlar neredeyse hakiki deneysel dizaynlarla aynıdır, fakat kilit bir öge hariç: rastgele atama. Mesela, okulun bir sınıfı veya bir organizasyon muamele grubu olarak kullanılırken, aynı okulun bir başka sınıfı veya aynı sektörden farklı bir organizasyon kontrol grubu olarak kullanılır. Bu rastgele atama yoksulluğu, potansiyel olarak harici faktörler bakımından denkleşmemiş grupların oluşmasıyla sonuçlanır, mesela bir grup belli bir vasıfça diğerinden üstün olabilir, diyelim okulun en iyi öğretmenine sahip olmak gibi. Bu da, *seçim yanlılığı*na olanaklı kılar. Yarı-deneysel dizaynlar, bu sebeple, hakiki deneysel dizaynlardan, içsel geçerlilik bakımından, daha aşağıdır. Çünkü seçimle alakalı çeşitli tehditler karşısında savunmasızdır. Mesela, seçim-maturasyon tehdidi (muamele ve kontrol gruplarının farklı hızlarda olgunlaşmaları), seçim-tarih tehdidi (muamele ve kontrol gruplarının, harici veya tarihsel olaylar tarafından farklıca etkilenmeleri), seçim-regresyon tehdidi (muamele ve kontrol gruplarının, öntest ve sontest arasındaki aritmetik-ortalamaya farklı hızlarda regresyonları), seçim-enstrümantasyon tehdidi (muamele ve kontrol gruplarının ölçüme farklıca yanıt vermeleri), seçim-test tehdidi (muamele ve kontrol gruplarının önteste farklıca yanıt vermeleri) ve seçim-mortalite tehdidi (muamele ve kontrol gruplarının diferansiyel eksilme oranlarına sahip olmaları). Bu seçim tehditlerine bakılırsa, mümkün olduğunca, yarı-deneysel dizaynlardan kaçınmak genellikle tercih edilebilir.

Çoğu hakiki deneysel dizayn, rastgele atamayı ihmal etmek suretiyle, yarı-deneysel dizayna dönüştürülebilir. Mesela, öntest-sontest kontrol grubu dizaynının yarı-deneysel versiyonu **eşdeğer-olmayan gruplar dizaynı (NEGD)** olarak adlandırılır ki Figür 10.8 de görüldüğü üzere, rastgele atama R rastgele-olmayan atama N ile yer değiştirir.

N	O	X	O
N	O		O

Figür 10.8. NEGD dizaynı

Yine, çapraz tekrarlama dizaynının yarı-deneysel versiyonu eşdeğer-olmayan çapraz tekrarlama dizaynı olarak adlandırılır (Figür 10.9 a bak).

N	O	X	O		O
N	O		O	X	O

Figür 10.9. Eşdeğer-olmayan çapraz tekrarlama dizaynı

İlaveten, hakiki deneysel dizaynlara tekabül etmeyen başka bazı yekta eşdeğer-olmayan dizaynlar da mevcuttur. Bu dizaynların nispeten daha kullanışlı olanlarını aşağıda tartışacağız.

Regresyon süreksizlik dizaynı (RDD). Bu bir eşdeğer-olmayan öntest-sontest dizaynıdır öyle ki denekler muamele ve kontrol grubuna, bir program ölçüsüne ait bir kesme skoruna göre atanırlar. Mesela, ağır hastalar yeni bir ilacın veya tedavi protokolünün etkililiğini test etmek için bir muamele grubuna atanabilir ve hafif hastalar kontrol grubuna tanabilir. Bir başka misal olarak, standartlaştırılmış sınav skorları düşük olan öğrenciler performanslarını yükseltmeyi hedefleyen bir iyileştirici müfredat programı için seçilebilirler ve standartlaştırılmış sınav skorları yüksek olan öğrenciler bu iyileştirici müfredat programı için seçilmezler. Dizayn notasyonu aşağıdaki üzere sunulabilir ki orada C kesme skorunu temsil ediyor:

C	O	X	O
C	O		O

Figür 10.10. RD dizaynı

Kesme skorunun kullanılmasından dolayı, gözlenen sonuçların, muameleden ziyade, kesme skorunun bir fonksiyonu olması olanaklıdır ki bu da içsel geçerlilik için yeni bir tehdittir. Mamafih, kesme skorunu kullanmak, bir yarı-deneysel muamelenin uygulanmasını temin etmesinin yanı sıra, aynı zamanda sınırlı veya maliyetli kaynakların, popülasyonda rastgele dağıtılmasından ziyade, daha fazla ihtiyacı olan insanlara dağıtılmasını da sağlar. Gruplar arasındaki sistematik eşedeğersizlik göz önüne alındığında, bir RD dizaynındaki kontrol grubu skorları, muamele grubu skorları ile karşılaştırılmak üzere bir kıstas olarak hizmet etmez. Daha ziade, eğer kontrol grubunda öntest ve sontest skorları arasında süreksizlik yoksa, fakat böyle bir süreksizlik muamele grubunda varsa, o zaman, bu süreksizlik, muamele etkisinin bir kanıtı görülür.

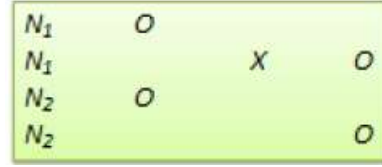
Proksi öntest dizaynı. Bu dizayn, Figür 10.11 de görüldüğü üzere, standart NEGD (öntest-sontest) dizaynına çok benzerdir; kritik bir farkla ki öntest skoru muamele uygulandıktan sonra toplanır. Bu dizaynın tipik bir uygulaması bir araştırmacının bir programın etkililiğini (mesela, bir öğretim programı), o program halihazırda başladıktan sonra ve öntest verisi mevcut değilken test etmeye kalkmasıdır. Böyle şartlar altında, araştırmacı için en iyi opsiyon, çoğu zaman, önceden kaydedilmiş farklı ölçüleri öntest verisi için bir proksi olarak kullanmaktır, mesela öğrencilerin programın başlamasından önceki not ortalamaları gibi. Proksi öntest dizaynının bir çeşidi deneklerin sontestte toplanmış öntest verisini kullanmaktır ki bu belki hatırlama yanlılığına açıktır fakat en azından yanıt değişkenindeki değişimin bir ölçüsünü temin edebilir.

N	O_1	X	O_2
N	O_1		O_2

Figür 10.11. Proksi öntest dizaynı

Ayrı öntest-sontest örnekleri dizaynı. Bu dizayn yararlıdır eğer ki öntest ve sontest verisini aynı deneklerden bir sebeple toplamak olanaksız ise. Figür 10.12 e gösterildiği üzere, bu dizaynda dört grup vardır, fakat iki grup tek bir eşdeğer-olmayan gruptan gelirken, diğer iki grup farklı bir eşdeğer-olmayan gruptan gelir. Mesela, bir şehirde sunulan yeni bir çevrimiçi hizmete dair müşteri memnuniyetini test etmek istiyoruz. Bu durumda, bu hizmetin sunulduğu şehirdeki müşteriler muamele grubunu oluştururken, bu hizmetin sunulmadığı başka bir şehirdeki müşteriler kontrol grubunu oluşturur. Eğer aynı müşterilerin öntest ve sontest puanlarını elde etmek olanaklı değilse, müşteri memnuniyetini zamanın bir noktasında ölçebiliriz, yeni hizmet programı tatbik edildikten sonra müşteri memnuniyetini farklı bir

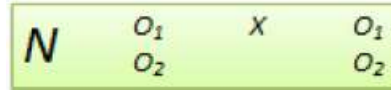
müşteri kümesi üzerinde ölçebiliriz. Müşteri memnuniyeti kontrol grubunda muamele grubu ile aynı zamanda ölçülmüş olur fakat bunlara yeni program tatbik edilmiştir. Bu dizayn bilhassa güçlü değildir çünkü herhangi bir spesifik müşterinin memnuniyet skorundaki değişmeyi (yeni hizmetin tatbikinden önceki ve sonraki) inceleyebilmeyiz fakat yalnızca birbirinden ayrı iki müşteri grubunun ortalama memnuniyet skorları arasındaki farkı inceleyebiliriz. Zayıf içsel geçerliliğine rağmen, aynı deneklerin öntest ve sontest verisi müsait olmadığında, bu dizayn hala yarı-deneyssel veri toplamının faydalı bir yolu olabilir.



Figür 10.12. Ayrı öntest-sontest örnekleri dizaynı

Eşdeğer-olmayan bağımlı değişken (NEDV) dizaynı. İki çıkım ölçüsüne sahip tek-grup ön-son yarı-deneyssel dizayndır. Çıkım ölçülerinden birinin teorik olarak muamelenin tesiri altında kalması ve diğer ölçünün kalmaması beklenir. Mesela, lise öğrencileri için yeni bir kalkülüs müfredatı dizayn ediyorsak, bu müfredat muhtemelen öğrencilerin sontest kalkülüs skorlarına tesir edecek fakat cebir skorlarına değil. Mamafih, mesela olgunlaşma gibi ekstra faktörlerden dolayı sontest cebir skorları yine de değişebilir. Bu sebeple, ön-son kalkülüs skorları muamele ölçüsü olarak alınırken, ön-son cebir skorları bir kontrol ölçüsü olarak kullanılabilir. Figür 10.13 te gösterilen dizayn notasyonu tek grubu tek N ile belirtir ve onu aynı grup öğrencileri için kalkülüs ve cebir için öntest O_1 ve sontest O_2 takibeder. Bu dizayn içsel geçerlilik bakımından zayıftır fakat ayrı bir kontrol grubu kullanmayı gerektirmemesi de avantajıdır.

NEDV dizaynının ilginç bir çeşidi *biçim eşleme NEDV dizaynıdır*, öyle ki bu dizayn çoklu çıkım değişkenlerini ve her değişkenin muameleden ne kadar etkileneceğini açıklayan bir teoriyi istihdam eder. Böylece, araştırmacı teorik tahminin fiili gözlemlerle eşleşip eşleşmediğini inceleyebilir. Teorik ve gözlenen biçimler arasındaki Bu *biçim-eşleme* tekniği orjinal NEDV dizaynındaki içsel geçerlilik endişelerini hafifletmenin kudretli bir yoludur.



Figür 10.13. NEDV dizaynı

Deneyssel Dizaynın Riskleri

Deneyssel dizayn araştırma dizaynlarının en zorlarından biridir ve hafife alınmamalıdır. Bu tip araştırma en iyisidir fakat yığınla metodolojik problemi de çoğu zaman beraberinde getirir. Birincisi, deneyssel araştırma test edilecek hipotezlerin formülasyonu için teorileri gerektirse bile, mevcut deneyssel araştırmanın çoğu teorik değildir. Teoriler olmadan, test edilen bir hipotez hususi (yani *ad hoc*), mantıksız ve manasız bir hipotez olma temayülü gösterir. İkincisi, deneyssel araştırmada kullanılan ölçüm enstrümanlarının pek çoğu güvenilirlik ve geçerlilik için test edilmez ve öğrenme-çalışmaları arasında karşılaştırılabilirlerdir. Vargısal olarak, böylesi enstrümanların kullanılmasıyla üretilen sonuçlar yine karşılaştırılabilirlerdir. Üçüncüsü, pek çok deneyssel araştırmada müteneşip olmayan araştırma dizaynları kullanılır, mesela münasebetsiz yanıt değişkenleri, interaksiyon etkilerinin ihmali, deneyssel kontrollerin ihmali ve muamele grupları arasında eşdeğer-olmayan dürtü gibi. Böylesi öğrenme-çalışmalarının bulguları içsel geçerlilik bakımından zayıflık eğilimi gösterirler ve çok şüphelilerdir.

Dördüncüsü, deneyisel arařtırmada kullanılan muameleler (görevler) öğrenme-çalışmaları arasında ihtilaflı, karşılaştırılabilir, ve tutarsız olabilirler ve bazen denek popülasyonu için münasebetsiz olabilirler. Mesela, lisans öğrencisi deneklere kendilerini bir satış müdürünün yerine koymaları ve kompleks bir bütçe aktarımı görevini icra etmeleri istenir, öyle ki bu öğrencilerin bu göreve dair hiç deneyimleri yoktur. Böyle mütenasip olmayan görevlerin kullanılması, içsel geçerliliğe karşı yeni tehditler arzeder (yani, denegin performansı görev kurulumunun içeriğı veya zorluğu ile paralel bir yapaylık yaratabilir), interpretasyonu yapılamaz ve manasız bulgular üretebilir ve farklı öğrenme-çalışmalarının bulgularının entegrasyonunu olanaksızlaştırabilir.

Münasip deneyisel muamelelerin dizaynı deney dizaynında çok önemli bir görevdir, çünkü muamele, deneyisel metodun varlık sebebidir ve asla aceleye getirilmemeli ve savsaklanmamalıdır. Kifayetli ve mütenasip bir görev dizayn etmek için, arařtırmacılar eğer mevcutsa önceden-geçerlenmiş görevleri kullanmalı, görevlerin kifayetini kontrol etmek üzere muamele manipölasyon kontrolleri yapmalı (tayin edilen görevi icra ettikten sonra, deneklerin ifadelerini almak suretiyle), pilot testleri yapmalı (eğer lüzumluysa, yineleyerek) ve eğer şüphe duyarlarsa yanıtlayıcılar örneğine kompleks veya yabancı gelen görevler yerine daha basit ve bilinen görevleri kullanmalılardır.

Hülasa, bu bölümde deneyisel dizayn arařtırma metodunun kilit kavramları taktim edildi ve hakiki deneyisel ve yarı-deneyisel dizaynların çeşitlerine değinildi. Bu dizaynlar her ne kadar içsel geçerlilik bakımından çok değışiklik gösterebilir, içsel geçerliliğı düşük dizaynlar yine de gözardı edilmemelidir çünkü spesifik şartlarda ve ampirik kontenjanslarda yararlı olabilirler.

Bölüm 11

Vaka Araştırması

Vaka öğrenme-çalışması olarak ta adlandırılan vaka araştırması bir fenomeni bir veya bir kaç mekanda kendi doğal ortamı (yani kurulumu/*setting*) içerisinde zaman sürecinde yoğun olarak öğrenmeye-çalışan bir metottur. Mesela görüşmeler, gözlemler, önceden kaydedilmiş dökümanlar ve ikincil veri gibi çoklu veri toplama metotları istihdam edilebilir ve ilgilenilen fenomen hakkında zengin, detaylı ve bağlamlştırılmış çıkarımlar temin edebilir. Vaka araştırması teori test etme maksadıyla pozitivist bir tavırla ve teori inşa etme maksadıyla interpretivist bir tavırla istihdam edilebilir. Bu metot iktisadi araştırmalarda diğer sosyal bilim disiplinlerine göre daha popülerdir.

Vaka araştırması mesela deneyler ve yoklama araştırması gibi rakip araştırma metotları karşısında kendine mahsus bazı güçlü hususiyetlere sahiptir. Birincisi, pozitivist metotlar yalnızca teori test etmede kullanılabilirken, vaka araştırması hem teori test etmede hem de teori inşa etmede kullanılabilir. İnterpretif vaka araştırmasında, ilgilenilen yapıların önceden bilinmesine ihtiyaç yoktur, araştırma ilerlerken bunlar veriden hasıl edilebilir. İkincisi, araştırma soruları araştırma prosesi esnasında modifiye edilebilir, eğer orjinal sorular daha az münasebetli veya abartılı bulunursa. Veri toplandıktan sonra, hiç bir pozitivist metotta bu olanaklı değildir. Üçüncüsü, vaka araştırması, bağlamsal veriyi daha geniş ve daha derin boyutlarıyla yakalama kabiliyeti sayesinde, diğer çoğu araştırma metoduna göre, ilgilenilen fenomenin daha zengin, daha bağlamlştırılmış ve daha otantik interpretasyonunun türetilmesine yardımcı olabilir. Dördüncüsü ilgilenilen fenomen çoklu katılımcıların perspektiflerinden ve analizin çoklu düzeyleri kullanılarak (mesela, bireysel ve organizasyonel) öğrenilmeye-çalışılabilir.

Aynı zamanda, vaka araştırması bazı içkin zayıflıklara sahiptir. Hiç bir deneysel kontrol ihtiva etmediği için, çıkarımların içsel geçerliliği her zaman zayıftır. Elbette, deneyler hariç, tüm araştırma metotları için bu müşterek bir problemdir. Mamafih, daha sonra betimleneceği üzere, vaka araştırmasında, “doğal kontroller” kullanılmak suretiyle, bu kontrol problemine cevap aranabilir. İkincisi, vaka araştırmasından türetilen neticelerin niteliği büyük ölçüde araştırmacının entegrasyon kabiliyetine bağlıdır. Deneyimli bir araştırmacı vaka verisinde acemi bir araştırmacının ıskalayabileceği kavramları ve biçimleri görebilir. Bu yüzden, bulgular bazen subjektif olmakla eleştirilir. Nihayetinde, çıkarımların ağır bir şekilde bağlamsallaştırılmış olmasından ötürü, vaka araştırmasından yapılacak çıkarımları başka bağlamlara veya organizasyonlara genelleştirmek zor olabilir.

Şunu anlamış olman önemli ki vaka araştırması mesela iktisat sınıflarında tartışılan Harvard vaka öğrenme-çalışmaları gibi *vaka betimlemelerinden* farklıdır. Sınıf tartışmasını ve öğrencilerin eleştirel düşünmesini stimüle etmek veya spesifik bir problemi bir organizasyonun nasıl ele aldığını analiz etmek amacıyla, vaka betimlemeleri tipik olarak bir organizasyonel problemi zengince detaylandırarak betimlerken, vaka araştırması formal bir araştırma tekniğidir ki organizasyonel fenomenlerin açıklamalarını türetmek için bir bilimsel metodu ihtiva eder.

Vaka araştırması zor bir araştırma metodudur öyle ki araştırmacının ileri araştırma becerilerine sahip olmasını gerektirir ve bu yüzden, çoğu zaman hataya meyillidir. Benbasat ve ark. (1987) vaka araştırmasında sıkça karşılaşılan beş problem betimlediler. Birincisi, pek çok vaka araştırması öğrenme-çalışması spesifik araştırma sorularına sahip olmadan başlar ve bu yüzden herhangi bir spesifik cevaba veya içgörüsü yüksek çıkarımlara sahip olmadan biter. İkincisi,

vaka mekanları çoğu zaman araştırma soruları ile uyumlarından daha ziyade erişilebilirliklerine ve elverişliliklerine göre seçilir ve bu yüzden ilgilenilen araştırma sorusunu kifayyetle cevaplayamazlar. Üçüncüsü, araştırmacılar toplanan veriyi çoğu zaman çoklu vasıtalar kullanmak suretiyle geçiremezler veya nirengilemezler ki bu da görüşülen yanlı deneklerin verdiği yanıtlara dayanan yanlı interpretasyonlara yol açar. Dördüncüsü, pek çok öğrenme-çalışması verinin nasıl toplandığı (mesela, hangi görüşme soruları kullanıldı, hangi dökümanlar incelendi, görüşülenlerin herbirinin organizasyonel pozisyonları neydi vs.) ve analiz edildiği hakkında öyle az detay verir ki çıkarımların güvenilirliği hakkında şüphe uyandırabilir. Beşincisi, bir boylamsal araştırma metodu olarak sahip olduğu güce rağmen, pek çok vaka öğrenme-çalışması bir fenomeni boylamsal olarak baştan sona izlemez ve böylece organizasyonel proseslere ve kendi doğası gereği temporal olan fenomenlere yalnızca bir kesitsel ve sınırlı bakışı temin eder.

Vaka Araştırmasında Kilit Kararlar

Bir vaka araştırması metodunu göz önüne alan bir araştırmacı bazı kilit kararları vermelidir. Birincisi, öğrenilmeye-çalışılan araştırma soruları için bu doğru metod mudur? Vaka araştırması metodu, formatif aşamalarda teori inşa edilen alanlardaki münasebetli yapıların keşfedilmesini ihtiva eden keşifleyici öğrenme-çalışmaları için, katılımcıların deneyimlerinin ve eylemlerin bağlamının kritik olduğu öğrenme-çalışmaları için ve faktörlerden veya nedenlerden (ne) daha ziyade kompleks temporal prosesleri (bir fenomenin niçini ve nasılı) anlamayı hedefleyen öğrenme-çalışmaları için bilhassa mütenasiptir. Mesela organizasyonel değişme ve geniş-ölçekli teknoloji tatbiki projeleri gibi çoklu katılımcıları ve olayların interaksyonlu dizilimlerini ihtiva eden kompleks organizasyonel prosesleri öğrenmeye-çalışmak için bu metod çok uygundur.

İkincisi, bir vaka öğrenme-çalışması için mütenasip analiz ünitesi nedir? Vaka araştırması çoklu analiz ünitelerini eş zamanlı olarak inceleyebildiği için, araştırmacı, bir fenomeni birey düzeyinde mi, grup düzeyinde mi, organizasyon düzeyinde mi, yoksa çoklu düzeylerde mi öğrenmeye-çalışacağına karar vermelidir. Mesela, grup kararı verme veya grup çalışması üzerine bir öğrenme-çalışması, tek düzeyli bir analizin temin edebileceğinden daha zengin bir anlayış türetmek için, mesela grup aktivitelerine bireysel katılım gibi birey düzeyindeki yapıları ve mesela grup kohezyonu ve grup liderliği gibi grup düzeyindeki yapıları biraraya getirebilir.

Üçüncüsü, araştırmacı bir tekli-vaka dizaynı mı yoksa bir çoklu-vaka dizaynı mı istihdam etmelidir? Eğer durum kendine has ve ekstrem ise, eğer açığa vurucu (yani, bilimsel araştırma için önceden erişilemeyen bir durum) ise veya iyi-formüle-edilmiş bir teoriyi test etmek için kritik veya karşıt bir vakayı temsil ediyor ise, tekli vaka dizaynı teori üretiminin başında daha mütenasiptir. Çıkarımların genelleştirilebilirliğini tesis etmek için ve fenomenin daha zengin ve daha nüanslı interpretasyonlarını geliştirmek için, teori test etmede, çoklu vaka dizaynı daha mütenasiptir. Yin (1984) çoklu vaka mekanlarının kullanılmasını, her bir vaka mekanını bir deneysel öğrenme-çalışmasına benzer kabul ederek ve pozitivist araştırmada kullanılabilecek katı bilimsel kurallara uyarak, tekrarlama mantığı ile tavsiye etmektedir.

Dördüncüsü, vaka araştırması için, hangi mekanlar seçilmelidir? Vaka araştırmasından türetilen çıkarımların bağlamsal mahiyetini göz önüne alırsak, mekan seçimi bilhassa kritiktir bir meseledir çünkü yanlış mekanı seçmek bizi yanlış çıkarımlara götürebilir. Eğer araştırmanın amacı teorileri test etmek veya çıkarımların genelleştirilebilirliğini incelemek ise, o zaman, gözlemlerin varyansını artırmak için, birbirine benzemeyen vaka mekanları seçilmelidir. Mesela, araştırmanın amacı firmalarda teknoloji tatbiki prosesini anlamak ise, teknoloji tatbiki prosesinin firma büyüklüğüne göre değişip değişmediğini incelemek için, büyük, orta ve küçük

firmaların bir karışımı seçilmelidir. Mekan seçiminde fırsatçılıktan veya elverişli olanı tercih etmekten daha ziyade araştırma sorularıyla uyuma önem verilmelidir ki burada “teorik örnekleme” adı verilen bir prosesi de anmak gerekir.

Beşincisi, vaka araştırmasında hangi veri toplama teknikleri kullanılmalıdır? Vaka araştırması için, her ne kadar görüşme (açık-uçlu/strüktüresiz veya odaklı/strüktürlü) açık ara en popüler veri toplama tekniği olsa bile, mesela doğrudan gözlem (mesela, yönetim toplantılarına, brifinglere veya planlama oturumlarına katılma), dökümantasyon (mesela, dahili raporlar, sunumlar ve andıçların yanı sıra gazete haberleri gibi harici raporlar), arşiv kayıtları (mesela, organizasyon şemaları, finansal kayıtlar vs.) ve fiziksel yapıntılar (mesela, aygıtlar, çıktılar, aletler) gibi başka teknikler ile görüşme verisi desteklenebilir. Dahası, araştırmacı görüşülenlerin yanıtlarını karşılaştırmak suretiyle, gözlenen veriyi nirengilemeli ve geçerlemelidir.

Vaka Araştırmasını Yürütmek

Vaka araştırması öğrenme-çalışmalarının en çoğu mahiyeti itibariyle interpretiftir. İnterpretif vaka araştırması öyle bir tümevarımsal tekniktir ki yeni teorilerin inşa edilmesi veya mevcut olanların genişletilmesi maksadıyla, kavramların ve biçimlerin ortaya çıkmasını temin etmek için, bir veya bir kaç vaka mekanında toplanan kanıtın sistematik olarak analizini ve sentezini ihtiva eder. Eisenhardt (1989)⁹ vaka araştırması ile teorilerin inşa edilmesine ait bir “yol haritası” teklif etmiştir. Bu yol haritasını aşağıda betimleyeceğiz. Pozitivist vaka araştırması için, aşağıdaki aşamaların bazılarını yeniden düzenlemeye veya modifiye etmeye ihtiyaç duyulabilir dahi; örnekleme, veri toplama ve veri-analitik teknikler genellikle aynı kalmaktadır.

Araştırma sorularının tanımlanması. Herhangi bir başka araştırma gibi, vaka araştırması da teorik ve pratik olarak ilginç araştırma sorularını tanımlayarak ve bu araştırma sorularına verilecek olanaklı cevaplar hakkında bazı sezgisel beklentileri veya ilk vaka dizaynına kılavuzluk edecek önsel yapılar yapıları teşhis ederek başlamalıdır. Pozitivist vaka araştırmasında, önsel yapılar teoriye dayanırken, interpretif araştırmada böyle teori ve hipotezler *ex ante* (yani öncesinde, varsayımsal olarak) telakki edilmez. Bu araştırma soruları ve yapıları, interpretif vaka araştırmasında, eğer ihtiyaç duyulursa, daha sonra değiştirilebilir fakat pozitivist vaka araştırmasında değiştirilmez.

Vaka mekanlarının seçilmesi. Araştırmacı vaka mekanlarını teşhis etmek için bir “teorik örnekleme” (rastgele örnekleme değil) prosesi kullanılmalıdır. Bu yaklaşımda, vaka mekanları istatistiksel mülahazalardan daha ziyade mesela, önceki vakaları tekrarlamak, önsel teorileri genişletmek veya teorik kategorileri veya kutupsal tipleri doldurmak gibi teorik mülahazalara dayanarak seçilir. Seçilen mekanların araştırma sorularının mahiyetine uygun olmasına, firma büyüklüğü, sektör, vesaire faktörlerin etkilerine matuf varyansın minimize edilmesini ve ilgilenilen açıklayıcı değişkenlerin etkilerine matuf varyansın maksimize edilmesini temin etmesine özen gösterilmelidir. Mesela, araştırmanın amacı bazı firmaların diğerlerine göre nasıl daha çok inovatif olduğunu incelemek ise; araştırmacı, sektör ve büyüklük etkilerini azaltmak için, aynı sektördeki benzer büyüklükteki firmaları seçmelidir ve firma inovasyonunun varyansını artırmak için, bazı daha çok inovatif firmaları ve bazı daha az inovatif firmaları seçmelidir. Potansiyel mekana kuru kuruya yazı yazmak veya telefon etmek yerine, her firmada, projeyi onaylayacak otoriteye sahip olan veya otorite sahibi bir kişiyi görevlendirebilecek, yönetici düzeyindeki bir kişi ile irtibat kurmak daha iyidir. İlk

⁹ Eisenhardt, K. M. (1989). “Building Theories from Case Research,” *Academy of Management Review* (14:4), 532-550.

konuşmalarda, araştırmacı projenin mahiyetini ve maksadını, vaka mekanına sağlayacağı herhangi potansiyel yaralarını, toplanacak verinin nasıl kullanılacağını, veri toplamada kimlerin (araştırmacı, asistanları, vs.) yer alacağını, görüşülmesi arzu edilen kişiler ve sponsor olacak organizasyonun harcaması gereken zaman, çaba ve para miktarını betimlemelidir. Araştırmacı aynı zamanda hem firma için hem de bireysel yanıtlayıcılar için güveni, mahremiyeti ve anonimliği garanti etmelidir.

Enstrümanların ve protokollerin yaratılması. Vaka araştırmasında veri toplamanın birincil modu görüşmeler olduğu için, görüşme prosesine kılavuzluk edecek bir görüşme protokolü dizayn edilmelidir. Bu, esasen, sorulacak soruların bir listesidir. Sorular açık-uçlu (strüktürsüz) olabilirler veya kapalı-uçlu (strüktürlü) olabilirler veya her ikisinin bir kombinasyonu olabilirler. Görüşme protokolü sıkı sıkıya takip edilmelidir ve görüşmeci görüşme prosesi esnasında soruların sırasınıdeğiştirmemeli veya herhangi bir soruyu atlamamalıdır. Bununla beraber, eğer yanıtlayanın yorumları müphem veya ilginç ise, daha derinlemesine irdellemek için, bazı sapmalara izin verilebilir. Görüşmeci nötr bir tonu sürdürmelidir, yanıtlayıcıyı herhangi bir spesifik yöne yöneltmemelidir, diyelim ki verdiği herhangi bir yanıtta, jest ve mimiklerle dahi olsa, katıldığını veya katılmadığını ifade ve ima etmemelidir. Daha detaylı görüşme teknikleri yoklamalara dair bölümde tartışılmaktadır. İlaveten, görüşme verisini desteklemek ve geçerlemek için, mesela dahili dökümanlar ve andıçlar, yıllık raporlar, finansal beyanlar, gazete makaleleri ve doğrudan gözlemler gibi ek veri kaynakları elde edilmeye çalışılmalıdır.

Yanıtlayıcıların seçilmesi. İlgilenilen fenomene ıraksak perspektiflerden bakabilmek için, görüşme yanıtlayıcılarını farklı organizasyonel düzeylerden, departmanlardan ve pozisyonlardan seç. Görüşülenlerin bir rastgele örnekleme en tercihe şayandır; bununla beraber, örnekte çeşitli perspektifler temsil edildiği sürece, bir *kartopu örneği* de kabul edilebilir. Görüşülenler, elverişliliklerine ve erişilebilirliklerine göre değil, soruşturulan fenomen ile olan kişisel münasebetlerine göre ve araştırmacının sorularını isabetlice ve kifayyetle yanıtlamaktaki kabiliyetlerine ve istekliliklerine göre seçilmelidir.

Veri toplamaya başlama. Daha sonra atıfta bulunmak için, görüşmeleri elektronik olarak kaydetmek çoğu zaman iyi bir fikirdir. Mamafih, böyle bir kayıt yalnızca görüşülenin rızası ile yapılmalıdır. Hatta görüşmeler kaydediliyor olsa bile, araştırmacı, önemli yorumları veya kritik gözlemleri, davranışsal yanıtları (mesela, yanıtlayıcının vücut dili) ve araştırmacının yanıtlayıcı ve onun yorumları hakkındaki kişisel izlenimlerini defterine yazmalıdır. Her görüşme tamamlandıktan sonra, görüşmenin tümü, analiz için, bir metin dökümanı olarak, kelimesi kelimesine yazıya geçirilmelidir.

Vaka-içi veri analizinin icra edilmesi. Veri analizi veri toplamayı takip edebilir veya onunla çakışabilir. Veri toplamayı ve analizi çakıştırmanın şu avantajı vardır ki veri analizinde ortaya çıkan temalara dayanarak, veri toplama prosesinde yeniden düzenleme yapılabilir veya bu temalar daha derinlemesine irdelenebilir. Veri analizi iki safhada yapılır. Birinci safhada (vaka-içi analizi), ilgilenilen probleme ait bir ilksel teori üretebilmek için, araştırmacı ayrı ayrı her vaka mekanındaki belirgin kavramları ve bu kavramlar arasındaki kalıpları incelemelidir. Araştırmacı, araştırma problemini manalandırmak için, vaka mekanındaki kendi kişisel gözlemleri veya deneyimleri ile birlikte, görüşme verisini subjektif olarak kullanabilir. Alternatif olarak; mesela Glasser ve Strauss'un (1967) grounded teori yaklaşımı gibi bir kodlama stratejisi benimsenmek suretiyle; bir kanıt ve çıkarımlar zinciri türetmek için, mesela açık kodlama, aksiyal kodlama ve seçici kodlama gibi teknikler kullanılabilir. Bu teknikler daha sonraki bölümde detaylıca tartışılacak. Yine, mesela verinin grafiksel temsili (mesela, ağ diyagramı) veya sekans analizi (boylamsal veri için) gibi ev-yapımı teknikler kullanılabilir. Şuna dikkatini çekmek isterim ki çeşitli tiplerdeki vaka verisini analiz etmenin önceden

tanımlanmış herhangi bir yolu yoktur ve veri-analitik teknikler, araştırma projesinin mahiyetine uygun olacak şekilde modifiye edilebilir.

Çapraz-vaka analizinin icra edilmesi. Çoklu-mekan vaka araştırması, veri analizinin ikinci safhası olarak, çapraz-vaka analizini gerektirir. Bu analizde, kendine özgü (*/idiosyncratic*) neticelere yol açabilecek bağlamsal farkları gözardı etmek suretiyle, araştırmacı, farklı vaka mekanları arasındaki benzer kavramları ve biçimleri aramalıdır. Böylesi biçimler ilksel teoriyi geçerlemede veya daha kapsayıcı ve genelleştirilebilir bir teori geliştirmek için, onu rafine etmede (kavramları ve ilişkileri ekleyerek veya çıkararak) kullanılabilirler. Bu analiz çeşitli formlarda yapılabilir. Mesela, araştırmacı kategorileri (mesela, firma büyüklüğü, sektör, vs.) seçebilir ve grup-içi bezerlikleri ve gruplar-arası farkları (mesela, performansı yüksek olanlara karşı performansı düşük olanlar, inovasyonculara karşı ağır kanlı olanlar) arayabilir. Alternatif olarak, firma çiftleri arasındaki benzerlikler ve farklar eşlemeli olarak listelenerek, firmalar karşılaştırılabilir.

Hipotezlerin kurulması ve test edilmesi. Vaka mekanları arasında genelleştirilebilir kavramlar ve temalar ortaya çıktığında, bunların üzerine, tentatif hipotezler inşa edilir. Gözlenen veriye uyup uymadıklarını görmek için, bu hipotezler gözlenen kanıt ile iteratif olarak karşılaştırılmalı; ve eğer değilse, yapılar veya ilişkiler rafine edilmelidir. Aynı zamanda, araştırmacı, ortaya çıkan yapıları ve hipotezleri, içsel geçerlilikleri ve genelleştirilebilirlikleri bakımından, önceki literatürde rapor edilenler ile karşılaştırmalıdır. Çatışan bulguları reddetmek yerine, nevezuhur teori hakkında daha fazla içgörü üretecek yaratıcı düşünme vasıtasıyla, uzlaştırma yoluna gidilmelidir. Teori ve veri arasındaki daha ileri iterasyonlar mevcut teoriye dair yeni içgörüler veya değişiklikler üretmediği zaman, “teorik doyuma” ulaşılmış ve teori inşası prosesi tamamlanmış olur.

Vaka araştırması raporunun yazılması. Raporu yazarken, araştırmacı; örneklemede, veri toplamada, veri analizinde ve hipotez geliştirmede kullandığı prosesleri detaylıca ve açıkça öyle betimlemelidir ki okuyucular rapor edilen çıkarımların makuliyetini, gücünü ve tutarlılığını bağımsızca değerlendirebilsinler. Bulguların, araştırmacının önyargıları tarafından yanlışlaştırılmadığını garanti etmek için, araştırma metotlarının yüksek bir düzeyde açıklığına ihtiyaç vardır.

İnterpretif Vaka Araştırması için Misal

İnterpretif vaka araştırmasını öğrenmenin belki de en iyi yolu nasıl yapıldığını gösteren bir misali incelemektir. Böylesi bir misal Eisenhardt’ın (1989)¹⁰ yapmış olduğu öğrenme-çalışmasıdır ki icracıların (*/executives*) yüksek-hızlı çevrelerde (HVE) nasıl karar verdikleri hakkındadır. Bu bölümdeki sinopsisi okumadan önce, *Acedemy of Management Journal*’da yayımlanmış olan orijinal makaleyi okumanı öğütlerim. Bu öğrenme-çalışmasında, Eisenhardt, diğer firmalardakiler hızlı karar alabilmezken, bazı HVE firmalardaki icracı takımların (*/executive teams*) hızlı kararları nasıl aldıklarını ve hızlı karar almanın böyle çevrelerde firma performansını artırıp artırmadığını veya düşürüp düşürmediğini inceledi. HVE öyle bir çevre olarak tanımlanır ki orada talep, rekabet ve teknoloji o kadar hızlı ve kesintisizce değişir ki mevcut enformasyon çoğu zaman isabetsizdir, namüsaittir veya eskimiştir. Zımnı varsayımlar şunlardır: (1) HVE’de kifayetsiz enformasyonla hızlı karar vermek zordur. (2) Hızlı kararlar etkin olmayabilir ve kötü firma performansına neden olabilir.

İcracı karar-verme (*/executive decisin-making*) üzerine literatürü gözden geçirdikten sonra,

¹⁰ Eisenhardt, K. M. (1989). “Making Fast Strategic Decisions in High-Velocity Environments,” *Academy of Management Journal* (32:3), 543-576.

Eisenhardt bazı biçimler bulmuş olsa da, bu biçimlerin hiç biri yüksek-hızlı çevreye spesifik değildir. Literatüre göre, günü kurtarmaya yönelik olarak, hızlı karar veren firmalar daha az kaynaktan girdi almakta, daha az alternatifi mülahaza etmekte, daha sınırlı analiz yapmakta, karar-vermede kullanıcı katılımını kısıtlamakta, karar-verme otoritesini merkezileştirmekte ve sınırlı içsel çatışmalar yaşamaktadır. Mamafih, Eisenhardt’a göre, bu görüşler karar vericilerin yüksek-hızlı çevrelerde nasıl karar verdiklerini illa ki açıklamayabilir, öyle ki böylesi çevrelerde kararlar çabukça ve eksik enformasyonla verilmelidir ve aynı zamanda verilen kararların kalitesi yüksek olmalıdır.

Bu fenomeni incelemek için, Eisenhardt, kişisel bilgisayar sektöründeki yedi firma ile bir tümevarımcı öğrenme-çalışması yürüttü. Bin dokuz yüz seksenlerde, UNIX işletim sisteminin, RISC mimarisinin ve 64KB rastgele erişimli belleğin piyasaya sunulması ile beraber, kişisel bilgisayar sektörü teknolojiye dramatik değişiklikler yaşıyordu. Aynı zamanda, IBM’in kişisel bilgisayar işine girmesi rekabeti artırmıştı. Yine, müşteri talebindeki iki haneli artış ile beraber, bu sektör profili yüksek-hızlı çevreye uymaktaydı. Tekrarlama mantığı güdülen, bu çoklu vaka dizaynında, her bir vakanın diğer vakalardan yapılan çıkarımları onaylamaları veya onaylamamaları beklenmektedir. Vaka mekanları, araştırmacının erişimine ve yakınlığına göre seçilmiştir. Ancak, seçilen firmaların tümü Kaliforniyadaki Silikon Vadisinde yüksek-hızlı bilgisayar sektöründe işletilen firmalardır. Firmaların aynı sektörden ve aynı bölgeden seçilmesi yanıt değişkenlerinde (karar hızı veya performans), sektöre veya coğrafi farklara atfedilebilecek “paraziti” veya varyansı izale etmiştir.

Öğrenme-çalışması, analizin çoklu-düzeylerini birarada barındıran bir dizaynı istihdam etmektedir: karar (her firma içerisinde çoklu stratejik kararları karşılaştırma), icracı takımlar (stratejik kararlardan sorumlu farklı takımları karşılaştırma) ve firma (genel firma performansı). Veri beş kaynaktan toplanmıştır:

- Baş icracılar (/Chief Executive Officers) ile ilk görüşmeler: CEO'lara, firmalarının rekabet stratejisi, ayrıcalık yaratan rekabet unsurları, başlıca rakipleri, performansları ve son verdikleri veya vermekte oldukları başlıca stratejik kararları hakkında sorular sorulmuştur. Bu görüşmelere dayanarak, daha ileri soruşturma için, her firmadan bir kaç stratejik karar seçilmiştir. Kararları seçerken, dört kriter kullanılmıştır: (1) Firmanın stratejik pozisyon alışına müdahil olan kararlar. (2) Yüksek maliyetli kararlar. (3) Çoklu fonksiyonlara müdahil olan kararlar. (4) O firmanın stratejik karar-verme prosesini temsil eden kararlar.
- Bölüm başkanları ile görüşmeler: Her bölüm başkanına onaltı açık-uçlu soru sorulmuştur. Bu sorular, firmalarının rekabet stratejisi, fonksiyonel stratejisi, üst yönetim takımının üyeleri, takım etkileşiminin frekansı ve doğası, tipik karar verme prosesi, seçilmiş olan kararların herbirinin nasıl verildiği ve o kararların verilmesinin ne kadar zaman aldığıdır. Görüşmeler 1.5 saatten 2 saate kadar sürmüştür ve bazıları 4 saate kadar uzatılmıştır. Yanıtlayıcıların algılarından veya yorumlarından daha ziyade olgulara ve fiili olaylara odaklanmak için, “mahkeme” tarzı bir soru sorma davranışı benimsenmiştir. Mesela, “bu ne zaman oldu”, “sen ne yaptın” gibi. Görüşmeler iki kişi tarafından yürütülmüştür ve veri olguların çapraz-kontrolü ve görüşmecinin ve not-alıcının izlenimleri ile geçerlenmiştir. Tüm görüşme verisi kaydedilmekle beraber, her görüşme esnasında, görüşmecinin görüşme sonundaki genel izlenimi de dahil olmak üzere, notlar alınmıştır. “Yirmi dört saat kuralı” uygulanarak, görüşmeden sonraki 24 saat içinde, detaylı saha notları tamamlanmıştır, öyle ki bazı veri veya izlenimler, anımsanmayarak, kaybedilmesin.
- Anketler: Her firmadaki icracı takım üyeleri bir yoklama anketi doldurmuşlardır ki bu anket kendi firmalarındaki çatışmayı ve güç dağılımını ölçmek için nicel veriyi elde etmeye matuftur.

- İkincil veri: Mesela icracı takımların (stratejik kararlardan sorumlu olan) demografikleri, firmaların finansal performansları gibi sektör raporları ve içsel dökümanlar incelenmiştir.
- Kişisel gözlem: Son olarak, araştırmacı örneğindeki iki firmanın 1-günlük strateji celselerine ve haftalık icracılar toplantılarına katılmıştır.

Veri analizi nicel ve nitel tekniklerin bir kombinasyonudur. Çatışmaya ve güce ait nicel veri, firmalar/kararlar arasındaki biçimler için analiz edilmiştir. Nitel görüşme verisi, birden çok icracı tarafından zikredilen profil vasıflarını (mesela, sabırsızlık) kullanmak suretiyle, karar iklim profilleri içerisinde kombine edilmiştir. Vaka-içi analizi için, kilit karar olaylarına ait icracı eylemlerini bir zaman çizelgesine yerleştirilerek, stratejik kararların her biri için karar öyküleri yaratılmıştır. Çapraz vaka analizi için, firma çiftleri, benzerlikleri ve farkları bakımından karşılaştırılmıştır, mesela karar hızı ve firma performansı gibi ilgilenilen değişkenlere göre kategorize edilmiştir. Bu analizlere dayanarak, firma kategorileri içindeki her bir karar öyküsünden tümevarımsal olarak tentatif yapılar ve önermeler türetilmiştir. Her karar vakası, önerilen ilişkileri onaylamak için, tekrar ziyaret edilmiştir. İncelenen farkları literatür ile uzlaştırmak ve vaka bulgularından yeni içgörüler üretmek için, çıkarılan önermeler mevcut literatürde yer alan bulgularla karşılaştırılmıştır. Nihayetinde, geçerlenen önermelerden yüksek-hızlı çevrelerdeki firmaların stratejik karar-vermelerine ait tümevarımsal bir teori sentezlenmiştir.

Bu çoklu vaka araştırmasından türetilen çıkarımlar, mevcut literatüre göre beklenen bazı karar-verme biçimlerine aykırı düşmüştür. Birincisi, yüksek-hızlı çevrelerdeki hızlı karar vericiler, önceki literatürün öne sürdüğü üzere daha az değil, aksine daha çok enformasyon kullanmışlardır. Bununla beraber, bu karar vericiler daha çok *gerçek-zamanlı* enformasyon kullanmışlardır (daha önceki literatürde yer almayan bir içgörü), ve böylece problemleri, fırsatları ve değişen şartları daha hızlı teşhis edip, cevap verebilmişlerdir. İkincisi, hızlı karar vericiler daha çok sayıda alternatifi incelemişlerdir, daha azını değil. Mamafih, yavaş karar vericiler daha az sayıda alternatifi *sırasıyla* incelerken, hızlı karar vericiler bu çoklu alternatifleri *eş-zamanlı olarak* mülahaza ettiler. Üçüncüsü, hızlı karar vericiler, literatürün öne sürdüğü üzere, karar vermeyi merkezileştirmediler veya başkalarının girdilerini kısıtlamadılar. Daha ziyade, bu firmalar iki-aşamalı bir karar prosesi kullanmışlardır ki birinci aşamada deneyimli danışmanlardan girdiler alınmış, ikinci aşamada hızlı bir karşılaştırma ve karar seçimi yapılmıştır. Dördüncüsü, literatürde beklendiği üzere, hızlı karar vericiler daha az çatışma yaşamamışlardır fakat, çatışmayı azaltmak ve karar-verme hızını artırmak için, daha iyi çatışma çözüm tekniklerini istihdam etmişlerdir. Sonuç olarak, hızlı karar vericiler, büyük kararların hızlıca verilmesini sağlayan, kendi kognitif, duygusal ve politik prosesleri sayesinde, üstün firma performansları göstermişlerdir.

Pozitivist Vaka Araştırması için Misal

Vaka araştırması, aynı zamanda, teorileri veya hipotezleri test etmek için, pozitivist bir tavırla da kullanılabilir. Böyle araştırmalar nadir olmakla beraber, Markus (1932)¹¹, Golden Triangle Company'de (uydurma bir ad) teknoloji tatbiki üzerine yaptığı bir öğrenme-çalışmasında misal kabilinden bir illüstrasyon temin etmiştir. Bu öğrenme-çalışmasının amacı, GTC'deki muhasebecilerin verimliliklerini ve performanslarını artırmak niyetiyle yeni tatbik edilen bir finansal enformasyon sisteminin niçin GTC'nin genel merkezindeki muhasepler tarafından desteklendiğini fakat GTC şubelerindeki muhasepler tarafından niçin FES'e direnç

¹¹ Markus, M. L. (1983). "Power, Politics, and MIS Implementation," *Communications of the ACM* (26:6), 430-444.

gösterildiğini öğrenmeye-çalışmaktır. İlgilenilen fenomenin yektağından dolayı, bu bir tekli-vaka araştırma dizaynıdır.

FES'e karşı gösterilen kullanıcı direncinin sebeplerini keşfetmek için, Markus üç alternatif açıklama ileri sürmüştür: (1) Sistem-belirlenim teorisi: sistemin kifayetsizliğine ilişkin mesela kendi teknik kusurları, zayıf ergonomik dizaynı veya kullanıcı dostu olmayışı gibi faktörler bu dirence neden olmuştur. (2) İnsan-belirlenim teorisi: muhasiplerin sistemi kullanmaya elverişli olmayan kognitif modları veya kişilik vasıfları gibi kullanıcılara ilişkin faktörler bu dirence neden olmuştur. (3) İnteraksiyon teorisi: Sisteme ve insanlara ilişkin faktörlerin yanı sıra bu iki faktör kümesi arasındaki interaksiyon bu dirence neden olmuştur. Spesifik olarak, interaksiyon teorisi şunu öne sürmektedir ki FES organizasyon-içi gücün yeniden dağıtılmasına yol açmıştır ve bundan faydalanan muhasipler sisteme destek olurken, FES tatbikinin bir sonucu olarak, organizasyonel statülerini, münasebetlerini veya güçlerini kaybeden muhasipler sisteme direnç göstermişlerdir.

Bu üç teoriyi test etmek için, Markus bu üç teorik açıklamanın her birinden beklenen alternatif çıkımları tahmin etmiştir ve bu tahminlerin GTC'de yaptığı gözlemler ile ne kadar eşleştiğini analiz etmiştir. Mesela, sistem-belirlenim teorisi kullanıcı direncinin nedeninin kifayetsiz bir sistem olduğunu öne sürdüğü için, sistemin teknik problemlerinin çözülmesi ile direnci ortadan kaldırılacağı düşünülmüştür. Müteakiben, FES sistemini çalıştıran bilgisayar daha güçlü bir işletim sistemi, çevrimiçi işlem (muhasabe enformasyonunu anında işlemeyi geciktiren başlangıçtaki yığın işlemekten) ve yöneticilerin yeni hesap açması için daha basitleştirilmiş bir yazılım ile yükseltilmiştir. Bu değişikliklerin yapılmasından bir yıl sonra, sisteme direnç gösteren kullanıcılar hala direnç göstermeye devam etmişlerdir ve değiştirileceğini ümit etmişlerdir.

İnsan-belirlenim teorisi direnç gösteren bireyleri direnç göstermeyenler ile değiştirmenin veya onlarla bereber çalıştırmanın bu bireylerin FES'e karşı dirençlerini azaltacağını tahmin etmiştir. Müteakiben, GTC bir görev rotasyonu ve hareketliliği politikası başlatarak, direnç gösteren şubelerdeki muhasipleri değiştirmiştir, fakat direnç devam etmekle kalmamış, bazı vakalarda artmıştır. Spesifik bir misal olarak, genel merkezde çalıştığı zaman, sistemin dizaynırlarından ve lehinde olan bir muhasip, şube kontrolörü görevine atandıktan sonra, sisteme direnç göstermeye başlamıştır. İnsan-belirlenim teorisine ait tahminlerin gerçekleşmesinin akamete uğraması, bu teorinin reddedilmesine yol açmıştır.

Son olarak, interaksiyon teorisi, tatbik-öncesi safhadaki gücün yeniden dağıtılmasını ve dengesizliğini ele almaksızın, ne sistemin ne de insanların değiştirilmesinin (yani, kullanıcıların eğitimi veya görev rotasyonu politikası) direnci azaltmayacağını tahmin etmiştir. FES'nin tatbikinden önce, GTC'deki şube muhasipleri kendi şubelerinin operasyonlarına ait tüm muhasabe verisine sahip olduklarını biliyorlardı. Bu veriyi kalın muhasabe defterlerinde tutuyorlardı, başkalarının bu veriye erişimini kontrol ediyorlardı ve periyodik raporların yayımlanmasından önce, sıradışı muhasabe olaylarını uyuşturabiliyorlardı. Genel merkez muhasipleri, merkezi raporlama ve konsolidasyon için ihtiyaç duydukları şube verisine erişmekte büyük ölçüde şube muhasiplerine bağımlıydı. FES, otomatik olarak, tüm veriyi kaynağından toplayarak, tek bir merkezi veritabanında konsolide ettiği için, şube muhasiplerine olan ihtiyacı ortadan kaldırmıştır, kendi şubelerine ait veri üzerindeki kontrollerini ve otonomilerini zayıflatmıştır ve yaptıkları görevi bir nevi münasebetsizleştirmiştir. Genel merkez muhasipleri, şimdi, şube muhasipleriyle muhatap olmaksızın, veri tabanını doğrudan sorgulayabilirler ve şubelerin verisine erişebilirler, şubelerin bireysel performanslarını analiz edebilirler ve birbirleriyle karşılaştırabilirler ve sıradışı biçimleri ve aktiviteleri icracı komitelere rapor edebilirler. Bu da, elbette, şubelerin güç kaybetmeleri ile sonuçlanır. Her ne kadar Markus bu teoriyi ampirik veri ile test etmemiş olsa dahi, onun organizasyonel gücün

dağıtılması hakkındaki gözlemleri, iki alternatif teorisinin reddedilmesi ile birarada düşünüldüğünde, intraksiyon teorisinin gerekçelendirilmesine götürmüştür.

Geleneksel Araştırma ile Karşılaştırma

Hipotez testini hedefleyen pozitivist vaka araştırması; kontrol altına alınmış gözlemlerin, kontrol altına alınmış çıkarımların (*/deductions*), tekrarlanabilirliğin ve bulguların genelleştirilebilirliğinin (yani, pozitivist araştırmanın geleneksel prensiplerinin) yokluğundan dolayı çoğu zaman doğa bilimi araştırmacıları tarafından eleştirilir. Mamafih, mütenasip vaka araştırması dizaynları ile bu eleştirilerin üstesinden gelinebilir. Mesela, kontrol altına alınmış gözlemler problemi vaka araştırmasında deneysel kontrolün veya istatistiksel kontrolün elde edilmesinin zorluğuna matuftur. Halbuki, vaka araştırmacıları, “doğal kontrolleri” istihdam etmek suretiyle, böylesi kontrol yokluklarını kompanse edebilirler. Markus’un öğrenme-çalışmasında bu doğal kontrol başlangıçta sistemin lehinde olan fakat şube kontrolörü olarak atandıktan sonra direnç göstermeye başlayan genel merkez muhasibiydi. Bu misalde, muhasibin davranışındaki değişiklik onun yeni şube görevine hamledilebilir. Mamafih, böyle doğal kontroller önceden bilinebilmez ve vaka araştırmacıları, eğer proaktif olarak böyle kontrolleri bulmaya çalışmıyorlarsa, onları gözden kaçırabilirler. Bu arada, mesela kuyruklu yıldızlar ve onların kompozisyonları hakkında çıkarımlar yapmak için, kuyruklu yıldızların dünyaya yeterince yakın geçmesini beklemek gibi doğal kontroller yine astronomi, jeoloji ve insan biyolojisi gibi doğa bilimi disiplinlerinde de kullanılır.

Kontrol altına alınmış çıkarım (*/deduction*) problemi çıkarımları destekleyecek yeterli nicel kanıtın yokluğuna matuftur, çünkü vaka araştırması verisi çoğunlukla nitel mahiyettedir. Hipotez testleri (mesela, t-testi) için nicel verinin yokluğuna rağmen, vaka araştırmasında, teorik mülahazalara dayanan davranışsal tahminler üretilerek ve bu tahminler zaman içerisinde test edilerek, kontrol altına alınmış çıkarımlar (*/deductions*) hala elde edilebilir. Markus bu stratejiyi kendi öğrenme-çalışmasında istihdam etmiştir, şöyle ki kullanıcı direnci için alternatif üç teorik hipotez üretmiş ve bu tahminlerin ikisi fiilen gözlenen davranış ile eşleşmediği zaman, onları reddetmiştir. Bu halde, hipotezler, matematiksel testler yerine, mantıksal önermeler kullanılarak test edilmiştir. Matematik mantığının bir alt kümesi olduğu için, yapılan bu hipotez testleri tam manasıyla istatistiksel çıkarımlar kadar geçerlidir.

Üçüncüsü, tekrarlanabilirlik problemi, belirli bir vaka mekanının yektalığı ve kendine özgülüğü (*/idiosyncrasy*) hasebiyle, aynı fenomeni tekrar gözlemenin zorluğuna matuftur. Mamafih, Markus’un üç teorisini bir illüstrasyon olarak kullanmak suretiyle, bir araştırmacı aynı teorileri başka bir veri mekanında test edebilir ki bu halde yeni vaka mekanının kendine özgü mahiyetine istinaden üç farklı tahmin ortaya çıkabilir ve bu tahminler benzer şekilde test edilebilir. Başka bir deyişle, her ne kadar vaka araştırmasının mekanı veya bağlamı tekrarlanabilir olmasa dahi, vaka araştırmasının çıkarımlarını tekrarlamak olanaklıdır.

Dördüncüsü, vaka araştırması başka çevrelere genelleştirilebilmeyen yekta ve tekralanabilmez fenomenleri incelemeye eğilimlidir. Doğa bilimlerinde genelleştirilebilirlik, ilave öğrenme-çalışmalarıyla tesis edilir. Benzerce, farklı bağlamlarda farklı tahminlerle yürütülen ilave vaka öğrenme-çalışmaları bulguların genelleştirilebilirliğini tesis edebilir, eğer ki böyle bulguların öğrenme-çalışmaları arasında tutarlı olduğu gözlenirse.

Son olarak, İngiliz filozof Karl Popper bilimsel teorilerin dört gereksinimini betimlemiştir: (1) Teoriler yanlışlanabilir olmalı. (2) Teoriler mantıksal olarak tutarlı olmalı. (3) Teoriler کافی derecede tahmin kabiliyetine sahip olmalı. (4) Teoriler rakip teorilere göre daha iyi bir açıklama temin etmeli. Vaka araştırmasında, ilk üç gereksinimin yerine getirilme düzeyleri, gözlenen bulguların serbestlik derecelerini artırmak suretiyle, artırılabilir. Mesela vaka mekanlarının

sayısını, alternatif tahminlerin sayısını ve incelenen analiz düzeylerinin sayısını artırmak, serbestlik derecesini artırmanın bazı yollarıdır. Markus, öğrenme-çalışmasında, çoklu grupların (şube muhasipleri ve genel merkez muhasipleri) davranışlarını inceleyerek ve çoklu (üç) rakip açıklamalar temin ederek bunu kotarmıştır. Popper'in dördüncü koşulu, bu öğrenme-çalışmasında, hipotezlerden birinin, diğer iki hipoteze göre, gözlenen kanıtla daha iyi eşleştiğı bulunduğı zaman kotarılmıştır.

Bölüm 12

İnterpretif Araştırma

Bir önceki bölümde, interpretif araştırmayı, veya daha spesifik olarak, interpretif vaka araştırmasını ele aldık. Bu bölümde, interpretif araştırmanın başka çeşitlerini gözden geçireceğiz. Anımsıyor olmalısın ki mesela laboratuvar deneyleri ve yoklama araştırması gibi pozitivist veya tümünden-gelimci metotlar spesifik olarak teori (veya hipotez) testine matufken, mesela eylem araştırması ve etnoğrafya gibi tüme-varımcı metotlar teori inşasına matuftur. Araştırmacının bir teori ile başlayıp, ampirik veri kullanarak, teorik postülatları test ettiği bir pozivist metodun aksine, interpretif metotlarda, araştırmacı veri ile başlar ve gözlenen veriden, ilgilendiği fenomen hakkında bir teori türetmeye çabalar.

“İnterpretif araştırma” terimi çoğu zaman “nitel araştırma” terimi ile aynı manada kullanılmakla beraber, iki kavram hayli farklıdır. İnterpretif araştırma bir araştırma paradigmasıdır (Bölüm 3’e bak) ki şu varsayıma dayanır: Sosyal gerçeklik tekil veya objektif değildir. Fakat, daha ziyade insan deneyimi ve sosyal bağlamlar tarafından şekillendirilir (ontoloji) ve bu yüzden de çeşitli katılımcılarının subjektif interpretasyonlarını uzlaştırmak suretiyle (epistemoloji), en iyi kendi sosyo-tarihsel bağlamı içerisinde öğrenilmeye-çalışılır. İnterpretif araştırmacılar sosyal gerçekliği kendi sosyal kurulumu içerisine gömülü ve o kurulumdan soyutlanmasını olanaksız addettikleri için, gerçekliğin “interpretasyonunu” bir hipotez testi prosesinden daha ziyade bir “manalandırma” prosesi olarak telakki ederler. Bu pozitivist veya fonksiyonelist paradigmaya terstir. Öyle ki pozitivist paradigma gerçekliğin bağlamdan nispeten bağımsız olduğunu, bağlamından soyutlanabileceğini ve mesela standartlaştırılmış ölçüler gibi objektif teknikler kullanılarak, bileşenlerine ayırmacı fonksiyonel bir tavırla, öğrenilmeye-çalışılabileceğini varsayar. Bir araştırmacının interpretif araştırmayı mı yoksa pozitivist araştırmayı mı tercih etmeli? Bu sorunun cevabı ilgilenilen fenomenin mahiyeti hakkındaki paradigmasal mülahazalara ve bu fenomenin en iyi hangi yoldan öğrenilebileceğine bağlıdır.

Mamafih, nitel araştırma ve nicel araştırma karşılaştırıldığı zaman, toplanacak verinin tipi ve verinin nasıl analiz edileceği hakkında ampirik veya veriye-dair mülahazalar söz konusudur. Nitel araştırma çoğunlukla mesela görüşmeler ve gözlemler gibi sayısal-olmayan veriye dayanırken, bunun aksine, nicel araştırma mesela skorlar ve metrikler gibi sayısal veriyi istihdam eder. Bundan ötürü, nitel araştırma mesela regresyon analizi gibi istatistiksel prosedürleri istihdam etmeye uygun olmamakla beraber, içerik analizi gibi teknikler kullanılarak kodlama yapılır. Bazen, kodlanmış nitel veri kodların frekansları cinsinden nicel olarak tablollaştırılır fakat bu veri istatistiksel analize tabi tutulmaz. Pek çok bağınaz interpretif araştırmacı bu kodlama yaklaşımını, esasen subjektif olan bir sosyal fenomende beyhude bir konsensüs veya objektiflik arama çabası olarak görerek, reddederler.

Her ne kadar interpretif araştırma büyük ölçüde nitel veriye dayanma eğiliminde olsa dahi, nicel veri ilgilenilen fenomenin anlaşılmasına nitel veriye göre daha fazla kesinlik ve açıklık kazandırabilir.

Mesela, Eisenhardt (1989), yüksek-hızlı firmalarda karar verme üzerine yaptığı interpretif öğrenme-çalışmasında (bir önceki bölümde vaka araştırmasında tartıştık), firmaların belli stratejik kararları ne kadar sürede verdiklerine (ki 1.5 aydan 18 aya gidiyordu) ve her karar için kaç alternatif kararı mülahaza ettiklerine ait ait sayısal veri topladı. Yanı sıra, yanıtlayıcıların

organizasyonel çatışma algılarını bir yoklama ile belirledi. Böylesi sayısal veri bu araştırmacının, yanıtlayıcıların subjektif algılarına bel bağlamadan, yüksek-hızlı karar veren firmaları düşük-hızlı karar verenlerden açıkça ayırdetmesine yardımcı oldu. Ki bu da düşük-hızlı firmalarda ve yüksek-hızlı firmalarda mülahaza edilen alternatif kararların sayısını ve bu firmalardaki çatışma düzeyini karşılaştırmalı olarak incelemesine olanak sağladı.

İnterpretif araştırmada ilgilenilen fenomene dair hem nitel veri toplanacaktır ve hem de pozitivist araştırmada olduğu üzere nicel veri toplanacaktır. Nitel ve nicel verinin ortak kullanımı, çoğu zaman, “karışık-modlu dizayn” olarak adlandırılır ki araştırmacıya yekta içgörülerin yolunu açabilir ve bilimsel camiada son zamanlarda hayli itibarlıdır.

İnterpretif araştırma kökenini antropolojiden, sosyolojiden, psikolojiden, dil-bilimden ve semiyotikten alır ve pozitivist tekniklerin geliştirilmesinden çok daha önce, 19 uncu yüzyılın başından beri mevcuttur. İnterpretif araştırmada istihdam edilen nitel veri toplamanın ve interpretasyon prosesinin subjektif mahiyetinden dolayı, pek çok pozitivist araştırmacı bu tür araştırmayı hatalı ve yanlı olarak görür. Mamafih, pek çok pozitivist tekniğin ilgi-çekici iç-görüler veya yeni bilgi üretmekteki başarısızlıkları, 1970 li yıllardan başlayarak, interpretif araştırmaya karşı yeniden bir ilgi uyandırmıştır. Ancak interpretif çıkarımların güvenilirliğini ve geçerliliğini garanti edecek sağın metotların ve katı kriterlerin arayışı devam etmektedir.

Pozitivist Araştırmadan Ayrılan Hususiyetleri

Yukarıda tartışılan ontolojik ve epistemolojik varsayımlardaki temel paradigmatik farklılıkların yanı sıra, interpretif araştırma ve pozitivist araştırma başka yönlerden de farklılardır. Birincisi, interpretif araştırma bir *teorik örnekleme* stratejisi istihdam eder, öyle ki öğrenme-çalışması mekanları, yanıtlayıcılar veya vakalar mesela öğrenilmeye-çalışılan fenomene uyup uymadıkları (mesela, sürdürülebilir uygulamalar yalnızca sürdürülebilir uygulamaları tatbik eden organizasyonlarda öğrenilmeye-çalışılabilir), öğrenme-çalışması için onları yektaca uygun hale hale getiren belli karakteristiklere sahip olup olmadıkları (mesela, firma inovasyonlarının saikleri üzerine bir öğrenme-çalışmasına bazı yüksek inovasyonlu ve bazı düşük inovasyonlu firmalar dahil edilmelidir ki bu firmalar arasındaki karşıtlık ortaya çıksın) ve benzeri gibi teorik mülahazalara istinaden seçilirler. Aksine, pozitivist araştırmada *rastgele örnekleme* (veya bu tekniğin bir varyasyonu) istihdam edilir, öyle ki genelleştirilebilirlik maksadıyla, vakalar bir popülasyondan rastgele seçilir. Bundan ötürü, pozitivist araştırmada kabul edilebilir addedilmeyen elverişlilik örnekleri ve dar örnekler, interpretif araştırmada, öğrenme-çalışmasının mahiyetine ve maksadına uygun oldukları sürece, kabul edilebilir addedilir.

İkincisi, araştırmacının rolü interpretif araştırmada kritiktir. Mesela etnografya, eylem araştırması ve katılımcı gözlem gibi bazı metotlarda araştırmacı sosyal fenomenin parçası olarak telakki edilir ve onun araştırma prosesindeki spesifik rolü ve müdahalesi veri analizi sırasında açıkça belirtilmelidir. Mesela vaka araştırması gibi başka metotlarda, veri toplama ve veri analizi prosesleri sırasında araştırmacı “nötr” veya yansız bir duruş takınmalı ve kendi kişisel yanlılıklarının ve önyargılarının interpretif araştırmadan türetilen subjektif çıkarımların doğasını lekelemediğinden emin olmalıdır. Pozitivist araştırmada, mamafih, araştırmacının araştırma bağlamının dışında ve araştırma bağlamından bağımsız olduğu mülahaza edilir ve veri toplama ve veri analizi prosedürlerini yanlılaştırmadığı varsayılır.

Üçüncüsü, interpretif araştırma, redüksiyonist ve izolasyonist olmaktan daha ziyade, holistik ve bağlamsaldır. İnterpretif interpretasyonlar; pozitivist araştırmada çokça istihdam edilen istatistiksel tekniklerin tersine, sosyal fenomene dahil olan katılımcıların perspektifinden dile, işaretlere ve manalara odaklanma eğilimindedir. İnterpretif araştırmada katılık, yapı

geçerliliğine ait istatistiksel kıstaslardan veya hipotez testinden daha ziyade, veri toplama ve veri analizi için sistematik ve şeffaf yaklaşımlar cinsinden telakki edilir.

Dördüncüsü, veri toplama ve veri analizi, interpretif araştırmada, eş-zamanlı ve iteratif olarak devam ettirilebilir. Mesela, araştırmacı bir görüşmeyi yapabilir ve bir sonraki görüşmeye geçmeden evvel, onu kodlayabilir. Eş-zamanlı analiz araştırmacının görüşme protokolündeki potansiyel kusurları gidermesine veya ilgilenilen fenomeni daha iyi yakalayacak şekilde onu düzeltmesine yardımcı olur. Eğer araştırmacı orijinal araştırma sorusunun yeni ve yararlı iç-görüler üretebilmeyeceğinin farkına varırsa, orijinal araştırma sorusunu dahi değiştirebilir. Bu, interpretif araştırmanın değerli fakat çoğu zaman hafife alınan faydalı bir hususiyetidir ve pozitivist araştırma için söz konusu bile değildir. Pozitivist araştırmada, bir defa veri toplamaya başlandıktan sonra, tüm projeyi baştan almaksızın, araştırma projesi modifiye edilemez veya değiştirilemez.

İnterpretif Araştırmanın Faydaları ve Zorlukları

İnterpretif araştırma çeşitli yekta avantajlara sahiptir. Birincisi, mesela firmalar-arası ilişkiler veya ofisler-arası politikalar gibi nicel kanıtın yanlı, isabetsiz veya elde edilmesi bir şekilde zor olabildiği kompleks, biribiriyle-ilişkili veya çok yönlü sosyal proseslerin arkasındaki gizli sebepleri keşfetmek için çok uygundur. İkincisi, a-priori teorinin hiç olmadığı veya yetersiz olduğu alanlarda teori inşa etmede çoğu zaman yararlıdır. Üçüncüsü, bağlam-spesifik, yekta veya kendine-özü olaylar veya prosesler için mütenasiptir. Dördüncüsü, interpretif araştırma, aynı zamanda, takip araştırması için ilginç ve münasebetli soruların ve meselelerin açığa çıkarılmasına yardımcı olabilir.

Öte taraftan, interpretif araştırmanın kendine mahsus zorlukları da var. Birincisi, bu tip araştırma, veri toplama ve analiz süreçlerinde, pozitivist araştırmaya göre, daha yoğun zaman ve kaynak gerektirir. Çok az veri yanlış veya prematüre varsayımlara yol açabilirken, çok fazla veri araştırmacı tarafından efektif olarak işlenemeyebilir. İkincisi, interpretif araştırma iyi-eğitilmiş araştırmacılar gerektirir. Öyle ki bunlar kompleks sosyal fenomenleri, gömülü katılımcıların perspektiflerinden görebilmeliler ve interpretasyonunu yapabilmelilerdir. Kendi kişisel yanlılıklarını veya peşin hükümlerini yapacakları çıkarımlara enjekte etmeksizin, bu katılımcıların muhtelif perspektiflerini telif edebilmelilerdir. Üçüncüsü, tüm katılımcılar veya veri kaynakları ilgilenilen fenomen hakkında eşit olarak güvenilir, yansız veya bilgi sahibi olamayabilirler veya açığa vurmadıkları politik ajandaları olabilir ki bunlar yanıltıcı veya yanlış izlenimlere yol açabilir. Katılımcılar ve araştırmacı arasında karşılıklı güvenin eksikliği, katılımcıların kendilerini tam olarak ve dürüstçe ortaya koymalarına mani olabilir. “Dumanlı havada görebilmek” (yani, gizli veya yanlı ajandaların üstesinden gelebilmek) ve problemin gerçek mahiyetini anlamak interpretif araştırmacının vazifesidir. Dördüncüsü, interpretif araştırmadan yapılan çıkarımların aşırı bağlamsallaştırılmış doğası gereğince, böyle çıkarımların tekrarlanabilirliği veya genelleştirilebilirliği çok ta aranmamalıdır. Beşincisi, interpretif araştırma ilgilenilen araştırma sorularını cevaplama veya gelecekteki davranışları tahmin etmede bazen başarısız olabilir.

İnterpretif Araştırmanın Karakteristikleri

Her interpretif araştırma, aşağıda betimlendiği üzere, müşterek bir prensipler kümesine merbut olmalıdır.

Doğacı tahkikat: Sosyal fenomenler kendi doğal kurulumları içerisinde öğrenilmeye-çalışılmalıdır. İnterpretif araştırma sosyal fenomenlerin kendi sosyal bağlamları içerisinde yerleşik olduklarını ve kendi bağlamlarından izole edilebilmeyeceklerini varsaydığı için, böyle fenomenlerin interpretasyonları kendi sosyo-tarihsel bağlamı içerisinde köklendirilmelidir. Bu

şu manaya geliyor ki ilgilenilen bir fenomenin açıklamalarını bulmaya çabalarken, her ne kadar bağlamsal duyarlılık çıkarımların genelleştirilebilirliğini sınırlandıracak olsa dahi, bağlamsal değişkenler gözlenmeli ve mülahaza edilmelidir.

Enstrüman olarak araştırmacı: Araştırmacılar çoğu zaman öğrenmeye-çalıştıkları sosyal bağlamanın içerisine gömülülerdir ve veri toplama enstrümanının bir parçası olarak telakki edilirler. Bu halde, araştırmacılar gözlemsel becerilerini, katılımcılar nezdinde güvenilirliklerini ve doğru enformasyona ulaşma kabiliyetlerini kullanmak zorudalardır. Dahası, araştırmacının kişisel iç-görüşü, bilgisi ve sosyal bağlamla olan deneyimi; ilgilenilen fenomenin doğru bir şekilde interpretasyonunu yapmak için, kritiktir. Aynı zamanda, araştırmacılar kendi kişisel yanlılıklarının ve peşin hükümlerinin tamamen farkında olmalı ve böylesi yanlılıkların fenomenin adil ve isabetli portresini ortaya koyabilmelerine mani olmasına izin vermemelilerdir.

İnterpretif analiz: Gözlemlerin interpretasyonu, sosyal bağlama gömülü katılımcıların gözünden yapılmalıdır. İnterpretasyon iki düzeyde gerçekleştirilmelidir. Birinci düzey, fenomeni sosyal katılımcıların subjektif perspektiflerinden görmeyi veya deneyimlemeyi ihtiva eder. İkinci düzey, ilgilenilen fenomenin bir “kalın betimlemesini (*/thick description*)” veya bir zengin hikaye anlatımını temin etmek için, katılımcıların deneyimlerinin manasını anlamaktır. Böylece, katılımcıların niçin öyle davrandıkları anlatılabilir.

Ekspresif dilin kullanılması: Katılımcıların sözel veya sözel-olmayan dilinin belgelenmesi ve bu dilin analizi interpretif araştırmanın entegre bileşenleridir. Öğrenme-çalışması hikayenin, bir makinenin değil de, bir kişinin gözünden görüldüğünü garanti etmelidir ve o kişinin duygularını ve deneyimlerini resmetmelidir ki okuyucular o kişiyi anlayabilsinler ve bağ kurabilsinler. İmgelemelerin, metaforların, istihzaların ve diğer konuşma figürlerinin kullanılması interpretif analizde çok yaygındır.

Zamansal doğası: İnterpretif araştırma çoğu zaman spesifik cevapları aramakla ilgilenmez, daha ziyade dinamik bir sosyal prosesi, zaman içerisinde o kendini açığa vururken, anlamakla veya “anlamlandırmakla” ilgilenir. Bundan ötürü, bu tür araştırma, ilgilenilen fenomenin evrimini bütünüyle yakalamak için, araştırmacının, uzatılmış bir zaman periyodunda, derinlemesine bir müdahalesini gerektirir.

Hermenetik çember: İnterpretif interpretasyon, gözlem parçaları (metin/text) ve sosyal fenomenin bütünlüğü (bağlam/context) arasında ileri ve geri hareket eden bir iteratif prosestir. Bu prosesin gayesi, onların görünürdeki uyumsuzluğunu uyumlaştırmak ve gömülü katılımcıların birbirinden farklı subjektif bakış açıları ve deneyimleri ile tutarlı bir teori inşa etmektir. Bir fenomenin anlamı/manası ve gözlemler arasındaki bu iterasyonlar, “teorik doyuma” ulaşıncaya kadar devam etmelidir. Öyle ki, burası herhangi bir ilave iterasyonun ilgilenilen fenomene dair daha fazla içgörü temin etmediği yerdir.

İnterpretif Veri Toplama

İnterpretif araştırmada veri çeşitli teknikler vasıtasıyla toplanır. En çok kullanılan teknik **görüşmedir** (yüz-yüze, telefonla veya odak grup). Yoklama araştırmasına dair bir önceki bölümde, görüşme tipleri ve stratejileri detaylıca tartışıldı. İkinci bir teknik **gözlemdir**. Gözlemsel teknikler doğrudan gözlem ve katılımcı gözlemdir. *Doğrudan gözlemde*, araştırmacı nötr ve pasif bir dış gözlemcidir ve ilgilenilen fenomene müdahalede bulunmaz (vaka araştırmasında olduğu gibi). *Katılımcı gözlemde*, araştırmacı fenomenin aktif bir katılımcısıdır ve araştırmacının girdileri veya salt mevcudiyeti öğrenilmeye-çalışılan fenomene tesir eder (eylem araştırmasında olduğu gibi). Üçüncü bir teknik **belgelemedir**. Belgelemede, ilgilenilen fenomene ait daha ileri içgöründe bulunmak veya başka destekleyici kanıt formları bulmak için;

mesela günlükler, elektronik mektuplar, yıllık raporlar finansal ifadeler, gazete makaleleri, web siteleri gibi dışsal ve içsel belgeler kullanılabilir.

İnterpretif Araştırma Dizaynları

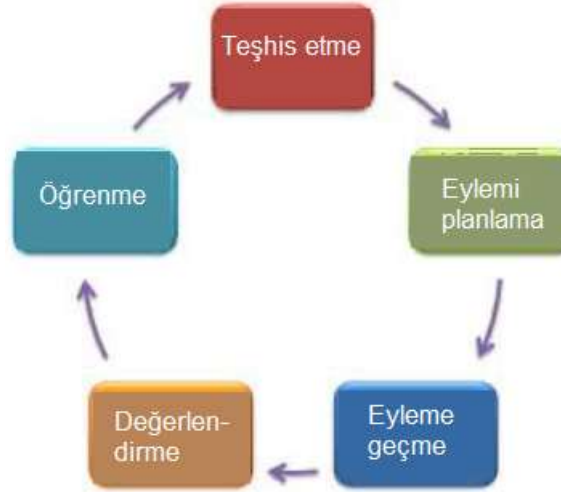
Vaka araştırması. Bir önceki bölümde tartışıldığı üzere, vaka araştırması, ilgilenilen bir fenomene ait detaylı bağlamsallaştırılmış çıkarımlar türetmek ve o fenomenin altında-yatan dinamik prosesi anlamak maksadıyla, bir fenomene ait bir veya birden çok araştırma mekanında yürütülen yoğun boylamsal bir öğrenme-çalışmasıdır. Vaka araştırması hem pozitivist bir yaklaşımla teori test etmede hem de interpretif bir yaklaşımla teori inşa etmede kullanılabilecek yegane araştırma dizaynıdır. Vaka araştırmasına dair bir önceki bölümde her iki teknik te derinlemesine tartışıldı ve aydınlatıcı misaller verildi. Dahası, vaka araştırmacısı sosyal kurulumda, aktif bir katılımcıdan (katılımcı gözlem) ziyade, nötr bir gözlemcidir (doğrudan gözlem). Herhangi başka bir interpretif yaklaşımda olduğu gibi, vaka araştırmasından manalı çıkarımlara varılması, araştırmacının gözlemsel becerilerine ve bütünleştirici yeteneklerine sıkıca bağlıdır.

Eylem araştırması. Eylem araştırması, teori inşasından ziyade teori test etmeye yönelik nitel fakat nicel bir araştırma dizaynıdır. (Bu bölümde ele alınmasının nedeni, münasip başka bir yer bulunmadığıdır.) Bu bir interaktif dizandır ki kompleks sosyal fenomenlerin, bu fenomenleri bir takım değişikliklere, müdahalelere veya “eylemlere” maruz bırakarak ve bu eylemlerin ilgilenilen fenomene ait çıkımlarını gözleyerek, en iyi anlaşılabilirliklerini varsayar. Bu metotta, araştırmacı çoğunlukla bir danışman veya bir sosyal bağlama (mesela bir organizasyon gibi) gömülü bir organizasyonel üyedir. Araştırmacı, bu rolde, bir sosyal probleme cevaben bir eylem başlatır ve eyleminin fenomene nasıl tesir ettiğini inceler. Aynı zamanda, eylem ve fenomen arasındaki ilişkiyi öğrenir ve bu ilişki hakkında içgörüler üretir. Eylemlere misal olarak; mesela bir organizasyonun performansının veya karlılığının kendi çalışma ortamı içerisinde artırılması amacıyla başlatılan yeni organizasyonel proseslerin, prosedürlerin, insanların veya teknolojilerin sunulması veya eskileriyle değiştirilmesi gibi organizasyonel değişim programları verilebilir. Araştırmacı eylem seçimini teoriye istinaden yapmalıdır ki seçtiği eylemlerin arzulanan sosyal değişimi niçin ve nasıl meydana getirebileceğini açıklayabilsin. Teori, seçilen eylemin hedeflenen problemi ne ölçüde çözdüğüyle, geçerlenir. Eş-zamanlı problem çözme ve içgörü üretimi, eylem araştırmasını diğer metotlardan (ki problem çözme ihtiva etmeyebilirler) ve danışmanlıktan (ki içgörü üretimini ihtiva etmiyor olabilir) ayıran başlıca hususiyetidir. Bundan ötürü, eylem araştırması, araştırma ve pratik arasında köprü kurmak için, mükemmel bir metottur.

Eylem araştırması metodunun çeşitli varyasyonları mevcuttur. Bunların en ünlü olanı Susan ve Evered (1978)¹² tarafından dizayn edilen katılımcı eylem araştırmasıdır. Bu metot beş safhadan oluşan bir eylem araştırması döngüsünü takip eder: (1) teşhis etme, (2) eylemi planlama, (3) eyleme geçme, (4) değerlendirme ve (5) öğrenme (Figür 11 e bak). *Teşhis etme* bir problemi kendi sosyal bağlamında tanımayı ve tanımlamayı ihtiva eder. *Eylemi planlama* problemin alternatif çözümlerini tanımayı ve değerlendirmeyi ve eylemin nasıl gerçekleştirileceğine (teorik rasyonele istinaden) karar vermeyi ihtiva eder. *Eyleme geçme* hazırlanan eylem planının tatbikidir. *Değerlendirme* safhası başlatılan eylemin orijinal problemi çözmede ne ölçüde başarılı olduğunu, yani, teorikleştirilen etkilerin sahiden pratikte gerçekleşip gerçekleşmediğini, inceler. *Öğrenme safhasında*, eylemin değerlendirilmesinden elde edilen deneyimler ve geri-bildirim, problem hakkında içgörüler üretmekte ve eylem için müstakbel modifikasyonlar veya iyileştirmeler öne sürmekte kullanılır. Değerlendirme ve öğrenme

¹² Susman, G. I. and Evered, R. D. (1978). “An Assessment of the Scientific Merits of Action Research,” *Administrative Science Quarterly*, (23), 582-603.

safhalarına istinaden, probleme daha iyi cevap verecek şekilde, eylem modifiye edilebilir veya düzeltilebilir ve eylem araştırması döngüsü modifiye edilen eylem dizisiyle yinelenir. Tüm araştırma döngüsünün en az iki defa üzerinden geçilmesi tavsiye edilir öyle ki birinci döngüden öğrenilen ikinci döngüde tatbik edilebilsin. Her ne kadar araştırmacının gözlemlerini teyit etmek için mesela görüşmeler ve belgesel kanıt gibi başka teknikler de kullanılmakla beraber, veri toplamanın birinci modu katılımcı gözlemdir.



Figür 10.1 Eylem araştırması döngüsü

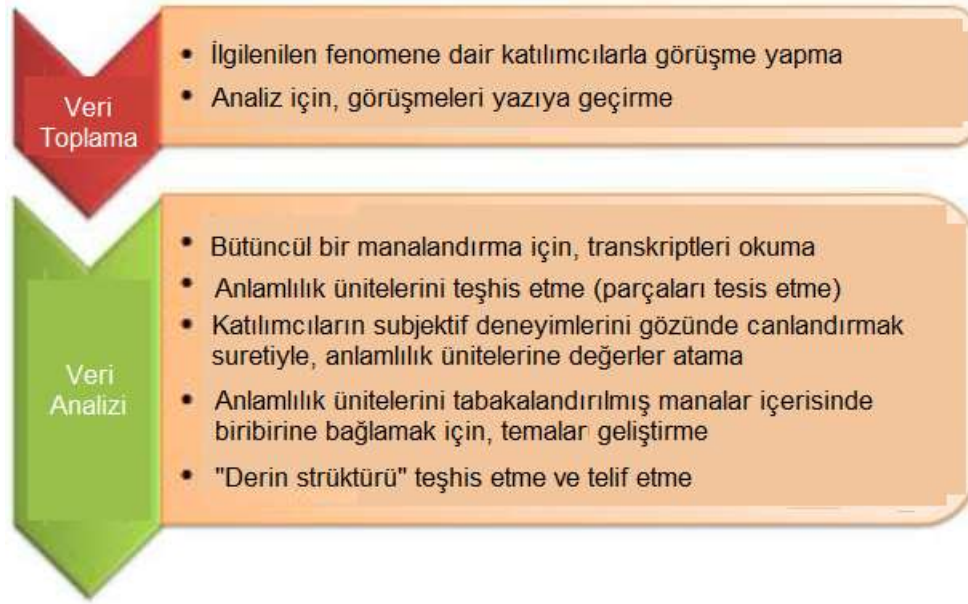
Etnografya. Etnografik araştırma metodu çoğunlukla antropoloji alanından türetilmiştir ve bir fenomeni kendi kültürü içerisinde öğrenmeye-çalışmayı öne çıkarır. Araştırmacı uzun bir zaman periyodu boyunca (çoğunlukla 8 aydan 2 yıla kadar) sosyal kültüre derinlemesine dalmalıdır ve öğrenilmeye-çalışılan kültürün ve onun katılımcılarının gündelik yaşantısına kendi doğal kurulumları içerisinde angaje olmalı, onları gözlemeli ve kaydetmelidir. Veri toplamanın birincil modu katılımcı gözlemdir ve veri analizi bir “anlama/manalandırma (/sense-making)” yaklaşımını ihtiva eder. İlaveten, araştırmacı alanda geniş kapsamlı notlar tutmalı ve deneyimini öyle detaylıca betimleyerek nakletmelidir ki okuyucular aynı kültürü araştırmacı kadar deneyimleyebilsinler. Bu metotta, araştırmacının iki rolü vardır: (1) kendi yekta bilgisine güvenmek ve içgörüler türetmeye angaje olmak (teori) ve (2) öğrenilmeye-çalışılan fenomenin geçişli-durumsal (/trans-situational) mahiyetine bilimsel camiayı ikna etmek.

Etnografik araştırmanın klasik misali Jane Goodall’ın primatlar üzerinde yaptığı öğrenme-çalışmasıdır. Bu öğrenme-çalışmasında, Jane Goodall Tanzanya’da Gombe Milli Parkı’nda kendi doğal habitatları içerisinde şempanzelerle beraber yaşadı; onların davranışlarını gözledi, onlarla etkileşti ve onların yaşantılarını paylaştı. Bu proses esnasında, şempanzelerin nasıl yiyecek ve barınak aradıklarını, birbirleriyle nasıl sosyalleştiklerini, iletişim biçimlerini, çiftleşme davranışlarını ve saire öğrendi. Etnografik araştırmanın daha çağdaş bir misali Myra Bluebond-Langer’in (1996)¹³ yaşamsal tehlikeyi havi hastalığı olan çocuklara sahip ailelerde karar verme ve karar vermeye tesir eden fiziksel, psikolojik, çevresel, etik, yasal ve kültürel meseleler üzerine yaptığı öğrenme çalışmasıdır. Bu araştırmacı tedavisi mümkün olmayan hastalıkları olan yaklaşık 80 çocuğun ve onların ailelerinin deneyimlerini iki yıldan uzun bir periyotta izledi. Veri toplama, katılımcı gözleminin yanı sıra, yaşadıkları deneyimi belgelemek için, çocuklarla, çocukların aileleriyle, çocukların yakınlarıyla ve çocuklara sağlık

¹³ Bluebond-Langer, M. (1996). *In the Shadow of Illness: Parents and Siblings of the Chronically Ill Child*. Princeton, NJ: Princeton University Press.

hizmeti verenlerle yapılan formal veya informal görüşmeleri ihtiva ediyordu.

Fenomenoloji. Fenomenoloji öyle bir araştırma metodudur ki içinde bulunduğumuz gerçekliği anlamının bir yolu olarak bilinçli deneyimleri öne çıkarır. Bu metot, tüm bilginin kaynağının insan deneyimi olduğuna inanan Alman filozofu Edmund Husserl’in 20 nci yüzyılın başlarında ileri sürdüğü fıkirlere dayanır. Fenomenoloji mesela insan yargısı, algıları ve eylemleri gibi bilinçli deneyimle irtibatlı fenomenlerin sistematik yansıması ve analizi ile ilgilenir. Bu metodun amacı (1) sosyal gerçekliği müdahil olan katılımcıların birbirinden farklı subjektif perspektiflerinden görmek ve betimlemek ve (2) bu subjektif deneyimlerin altında-yatan sembolik manaları (“derin strüktürü”) anlamaktır. Öğrenilmeye-çalışılan katılımcıları bilinçli düşüncelere, hislere ve davranışlara sevkeden derin strüktürleri tamamen anlayabilmek için, fenomenolojik tahkikat, araştırmacıların herhangi önsel varsayımlarını ve kişisel yanlılıklarını elemine etmelerini, katılımcının durumu ile empati kurlmalarını ve durumun varoluşsal boyutlarına inmelerini gerektirir.



Figür 10.2 Varoluşsal fenomenolojik araştırma metodu

Bazı araştırmacılar fenomenolojiiyi bir araştırma metodundan ziyade bir felsefe olarak görürler. Bu eleştiriye cevaben, Giorgi ve Giorgi (2003)¹⁴ bu alandaki öğrenme-çalışmalarına rehberlik etmek için, bir *varoluşsal fenomenolojik araştırma metodu* geliştirdi. Figür 10.2 de gösterilen bu metot veri toplama ve veri analizi safhalarına bölünebilir. *Veri toplama* safhasında, soruşturulan fenomene ilişkin deneyimlerini ve perspektiflerini yakalamak için, o sosyal fenomene gömülü katılımcılar ile görüşülür. Sorulara misal olarak şunları verebiliriz: “Tipik bir günü betimleyebilir misin?” “O belirli olayı daha detaylıca betimleyebilir misin?” Bu görüşmeler sesli kaydedilir ve ileride analiz edilmek üzere yazıya geçirilir. *Veri analizi* safhasında, (1) bütünsel bir bakışa sahip olmak ve (2) katılımcıların subjektif deneyimlerini doğrulukla temsil edebilecek “anlamlılık üniteleri” tesis etmek için, araştırmacı transkriptleri okur. Bu türden anlamlılık üniteleri için mesela “hissedilen-uzay”, “hissedilen-zaman” gibi kavramlar misal olarak verilebilir. Ki bunlar daha sonra katılımcıların psikolojik deneyimlerini belgelemek için kullanılır. Mesela, katılımcılar bir fenomeni deneyimlerken, kendilerini güvende, özgür, kısırılmış veya neşeli hissettiler mi (“hissedilen-uzay”) ? Deneyimlerinin

¹⁴ Giorgi, A. and Giorgi B. (2003). Phenomenology. In J A Smith (ed.) *Qualitative Psychology: A Practical Guide to Research Methods*. London: Sage Publications.

gergin, yavaş veya süreksiz olduğunu hissettiler mi (“hissedilen-zaman”)? Fenomenolojik analiz katılımcıların zamansal manzaralarını (yani, onların geçmiş, şimdi ve gelecek duyumlarını) hesaba katmalıdır ve araştırmacı kendini imgesel manada katılımcının durumuna aktarmalıdır (yani, katılımcının yaşamını zamansal olarak yaşamalıdır). Katılımcının yaşadığı deneyim anlatımsal bir formda veya ortaya-çıkan temalar kullanılarak betimlenir. O zaman, deneklerin yaşadığı deneyimlerin kırılğanlığını ve müphemliğini unutmaksızın, analiz, mananın çoklu katmanlarını teşhis etmek için, bu temaların içerilerine dalar.

İnterpretif Araştırmada Katılık (/Rigor)

Pozitivist araştırma, sosyal gerçekliği parsimonus teorilerle ve yasalarla basitleştirmek suretiyle, “indirgemeci” bir yaklaşımı benimserken, interpretif araştırma sosyal gerçekliğin interpretasyonunu, gerçekliğin yerleşik olduğu bağlamın içerisine gömülü katılımcıların subjektif bakış açıları üzerinden yapmaya girişir. Bu interpretasyonlar ağır bir şekilde bağlamsallaştırılır ve başka bağlamlara genelleştirilebilirlikleri elbette daha azdır. Mamafih, subjektifliği ve gömülü araştırmacının deneyimlerine ve içgörüsüne duyarlılığı nedeniyle, interpretif araştırma pek çok pozitivist (fonksiyonalist) araştırmacı tarafından daha az katı (/rigorous) bulunur. Pozitivist araştırmaya kıyasla, interpretif araştırma sosyal fenomen hakkında farklı ontolojik ve epistemolojik varsayımlar kümesine dayandığı için, mesela güvenilirlik, içsel geçerlilik ve genelleştirilebilirlik gibi katılığın pozitivist mefhumları aynı şekilde uygulanmazlar. Mamafih, Lincoln ve Guba (1985)¹⁵ interpretif araştırmanın katılığını (/rigor) yargılamak için kullanılabilecek alternatif bir kriterler kümesi temin etmiştir.

Emniyet (/Dependability). Eğer aynı fenomeni aynı kanıt kümesiyle değerlendiren iki araştırmacı bağımsızca aynı neticelere varırlarsa veya aynı veya benzer bir fenomeni farklı zamanlarda gözleyen aynı araştırmacı benzer neticelere varırsa, interpretif araştırmanın emniyetli (/dependable) veya otantik olduğu kabul edilebilir. Bu kavram pozitivist araştırmadaki güvenilirlik (/reliability) kavramı ile benzerdir. Öyle ki iki bağımsız araştırmacı arasındaki anlaşma, reytingciler-arası güvenilirlik mefhumu ile benzerdir ve aynı fenomenin aynı araştırmacı tarafından yapılan iki gözlemi arasındaki anlaşma test-yeniden-test güvenilirliği ile yakındır. Emniyeti garantilemek için, interpretif araştırmacılar ilgilendikleri fenomen ve o fenomenin içine gömülü olduğu sosyal bağlam hakkındaki ayrıntıları kifayyetlice temin etmelidirler öyle ki okuyucular kendi interpretif çıkarımlarını bağımsızca tasdik edebilsinler.

İtibar (/Credibility). Eğer okuyucular yapılan çıkarımları inanılır buluyorsa, interpretif araştırmanın itibarlı olduğu telakki edilebilir. Bu kavram fonksiyonelistik araştırmadaki içsel geçerlilik kavramı ile çok yakındır. Araştırmacının sahaya uzun süreli angajmanının kanıtı temin edilmek suretiyle, interpretif araştırmanın itibarı artırılabilir. Bunun için, denekler üzerinde veri nirengilemenin veya veri toplama tekniklerinin demonstrasyonu yapılabilir. Mesela görüşmelerin kelimesi-kelimesine transkripsiyonu, temasların ve görüşmelerin hassas kayıtları ve teorik ve metodolojik kararlara ait açıkça tutulan notlar gibi titiz veri idaresi prosedürleri ve analitik prosedürler istihdam edilebilir. Böylece, eğer gerekirse, bağımsız bir veri toplama denetimine olanak sağlanmış olur.

Onaylanabilirlik. Onaylanabilirlik interpretif araştırmanın rapor edilen bulgularının diğerleri tarafından (tipik olarak, katılımcılar) ne ölçüde bağımsızca onaylanabildiğine matuftur. Bu, fonksiyonelistik araştırmada objektiflik mefhumuna benzerdir. İnterpretif araştırma objektif bir gerçeklik mefhumunu reddettiği için, onaylanabilirlik “inter-subjektivite” terimiyle ifade edilir.

¹⁵ Lincoln Y. S., and Guba, E. G. (1985). *Naturalistic Inquiry*. Beverly Hills, CA: Sage Publications.

İnter-subjektivite, öğrenme-çalışmasının katılımcılarının araştırmacı tarafından türetilen çıkarımlarla aynı fikirde olup olmadıklarını yansıtır. Mesela, eğer bir öğrenme-çalışmasının katılımcıları ilgilenilen fenomen hakkında araştırmacı tarafından yapılan çıkarımlarla genellikle aynı fikirdeyseler, o zaman bulguların onaylanabilir olduğu telakki edilebilir.

Transfer-edilebilirlik. İnterpretif araştırmada transfer-edilebilirlik bulguların diğer kurulumlara (*/settings*) ne ölçüde genelleştirilebildiğine matuftur. Bu fikir fonksiyonelistik araştırmadaki dışsal geçerlilik fikrine benzerdir. Araştırmacı araştırma bağlamının zengin ve ayrıntılı betimlemelerini (“kalın betimleme”) temin etmeli ve veriden açığa çıkarılan yapıları, varsayımları ve prosesleri tafsilatıyla betimlemelidir. Öyle ki okuyucular rapor edilen bulguların diğer kurulumlara transfer-edilebilir olup olmadığını ve ne ölçüde transfer edilebileceğini bağımsızca değerlendirebilsinler.

Bölüm 13

Nitel Analiz

Nitel analiz mesela görüşme transkriptlerinde yer alan metin verisi gibi nitel verinin analizidir. İstatistiksel tekniklerin kullanıldığı ve araştırmacıdan büyük ölçüde bağımsız olan nicel analizin aksine, nitel analiz araştırmacının analitik ve bütünleştirici (/integrative) becerilerine ve verinin toplandığı sosyal bağlama dair kişisel bilgisine dayanır. Nitel analizde, bir fenomeni tahmin etmekten veya açıklamaktan ziyade, onu “manalandırmaya (/sense-making)” veya anlamaya önem verilir. Nitel araştırma için yaratıcı ve soruşturmacı bir zihinsel yapıya ihtiyaç vardır ki bu da etik olarak aydınlanmış ve bağlamda-katılımcı (/participant-in-context) bir tavır ve analitik stratejiler kümesine istinat eder. Bu bölümde, bu nitel analiz stratejilerinin bazıları kısaca gözden geçirilecek. Daha ilerisiyle ilgilenen okuyucu mesela Miles ve Huberman’ın (1984)¹⁶ bu konudaki seminal kitabı gibi daha otoriter ve daha detaylı referanslara başvurabilir.

Grounded Teori

Katılımcı gözlemi, derinlikli görüşmeler, odak gruplar, ses/görüntü kayıtlarındaki anlatımlar veya ikincil belgeler vasıtasıyla elde edilen devasa bir nitel veri kümesini nasıl analiz edebilirsin? Metin verisini analiz etmeye yarayan tekniklerden biri **grounded teoridir**. Bir sosyal fenomen hakkında teoriler inşa etmek için, o fenomene ait kaydedilmiş verinin interpretasyonunu yapan tümevarımcı bir tekniktir. Bu teknik Glaser ve Strauss (1967)¹⁷ tarafından geliştirildi ve daha sonra Strauss ve Corbin (1990)¹⁸ tarafından, spesifik kodlama tekniklerini de ihtiva edecek şekilde, rafine edildi. Kodlama, metin verisi segmentlerini bir kodlar kümesi (kavramlar), kategoriler (yapılar) ve ilişkiler şeklinde sınıflandıran ve kategorize eden bir prosestir. İnterpretasyonlar gözlenen ampirik veriye “oturtulur (/grounded)” (veya dayandırılır). Teorinin yalnızca gözlenmiş kanıta dayandığından emin olmak için, grounded teori yaklaşımı araştırmacının önceden var olabilecek teorik beklentilerini veya yanlılıklarını askıya almasını ve eldeki verinin teorinin formülasyonunu dikte etmesine izin vermesini gerektirir.

Strauss ve Corbin (1998), metin verisini analiz etmek için, üç kodlama tekniği betimlemektedir: açık, eksensel ve seçici. **Açık kodlama** öyle bir prosestir ki metinsel veri içerisinde saklı olan ve ilgilenilen fenomen ile potansiyel olarak ilişkili olan **kavramları** veya kilit fikirleri teşhis etmeyi hedefler. Araştırmacı, kavramlar olarak kodlanan ayrık olayları (/discrete events), hadiseleri (/incidents), fikirleri, eylemleri, algıları ve münasebetli interaksiyonları teşhis etmek için, ham metinsel veriyi satır satır inceler. Bunlar *in vivo kodlar* olarak adlandırılır. Her kavram, daha sonra geçerlenmek üzere, metnin spesifik parçalarıyla (kodlama üniteleri) irtibatlandırılır. Bazı kavramlar basit, açık ve müphematsız olabilirken diğerleri kompleks, müphem olabilirler ve farklı katılımcılar tarafından farklıca görülebilirler. Kodlama ünitesi çıkartılan kavramlar ile değişebilir. Mesela “organizasyon büyüklüğü” gibi basit kavramlar

¹⁶ Miles, M. B. and Huberman, A. M. (1984). *Qualitative Data Analysis: A Sourcebook of New Methods*. Newbury Park, CA: Sage Publications.

¹⁷ Glaser, B. and Strauss, A. (1967). *The Discovery of Grounded Theory: Strategies for Qualitative Research*, Chicago: Aldine.

¹⁸ Strauss, A. and Corbin, J. (1990). *Basics of Qualitative Research: Grounded Theory Procedures and Techniques*. Beverly Hills, CA: Sage Publications.

metnin yalnızca bir kaç kelimesini haviyken, mesela “organizasyonun misyonu” gibi kompleks olanları bir kaç sayfayı içerebilir. Kavramlar araştırmacının kendi adlandırma alışkanlıkları veya araştırma literatüründen alınan standartlaştırılmış etiketler kullanılarak adlandırılabilirler. Temel bir kavramlar kümesi, bir kere teşhis edildikten sonra, bu kavramlar verinin geri kalanını kodlamak için kullanılabilirler. Aynı zamanda, yeni kavramlar bulunmaya çalışılır ve eski kavramlar rafine edilir. Kodlama yaparken, her kavramın mesela büyüklüğü, rengi veya düzeyi gibi tanınabilir karakteristiklerini teşhis etmek önemlidir. Öyle ki benzer kavramlar daha sonra beraber gruplandırılabilir. Araştırmacının ilgilenilen fenomen ile münasebeti olan yeni kavramlara açık olmasından ve onları aktif olarak bulmaya çalışmasından dolayı, bu kodlama tekniği “açık” olarak adlandırılır.

Daha sonra, benzer kavramlar daha yüksek mertebeli **kategoriler** içerisinde gruplandırılırlar. Bir grounded teoride, kavramlar bağlam-spesifik olabilirken, kategoriler geniş ve genelleştirilebilir olma ve nihayetinde *yapılara* evrilme eğilimindedirler. Araştırmacının çalışmak zorunda olduğu kavramların miktarını azaltmak için ve bir sosyal fenomenin anlaşılmasında önem arzeden meselelerin bir “büyük resmini” çizmek için kategorilere ihtiyaç duyulur. Kategorileştirme iki safhada yapılabilir. Önce, kavramlar alt-kategorilerin içerisinde biraraya getirilir. Daha sonra, alt-kategoriler daha yüksek mertebeli kategoriler içerisinde biraraya getirilir. Bilhassa, eğer araştırmanın amacı halihazırdaki teorilerin genişletilmesi ise, bu kategorileri adlandırmak için mevcut literatürde bulunan yapılar kullanılabilir. Mamafih, mevcut yapıları kullanırken dikkatli olmak gerekir. Çünkü, mevcut yapılar kendileriyle beraber yaygınlaşmış inançları ve yanlışlıkları da getirebilirler. Her bir kategorinin karakteristikleri (veya sahip olduğu özellikler) ve her bir karakteristiğin boyutları teşhis edilmelidir. *Boyut* bir karakteristiğin bir süreklilik (*/continuum*) boyunca aldığı değeri temsil eder. Mesela, bir “iletişim ortamları” kategorisi “sürat” adı verilen bir karakteristiğe sahip olabilir, ki “hızlı, orta ve yavaş” olarak boyutlandırılabilir. Bu şekilde kategorileştirme iletişim ortamlarının farklı çeşitleri arasında ayırım yapmaya yardımcı olur ve araştırmacıların, mesela hangi iletişim ortamının hangi tip görevler için kullanıldığı gibi, verideki biçimleri (*/patterns*) teşhis etmelerine olanak sağlar.

Grounded teorisinin ikinci safhası **eksensel kodlamadır**, ki bu safhada kategoriler ve alt-kategoriler ilgilenilen fenomeni tentatif olarak açıklayabilen nedensel ilişkiler veya hipotezler içerisinde birleştirilir. Açık kodlamadan ayrı olmakla beraber, eksensel kodlama açık kodlamayla eş-zamanlı olarak yapılabilir. Kategoriler arasındaki ilişkiler veride açıkça görünür olabilirler veya daha ince ve zımni olabilirler. Sonki halde, araştırmacılar, hangi kategorilerin *koşulları* (fenomenin içine gömülü olduğu şartlar), *eylemleri/interaksiyonları* (bu koşullar altında bireylerin olaylara verdikleri yanıtlar) ve *vargıları* (*/consequences*) (eylemlerin/interaksiyonların çıkımları) temsil ettiklerini anlamak için bir kodlama şeması (ki çoğu zaman “kodlama paradigması” olarak adlandırılır, fakat Bölüm 3 te tartışılan paradigmalardan farklı) kullanabilirler. Koşullar, eylemler/interaksiyonlar ve vargılar teşhis edildikçe, teorik önermeler açığa çıkmaya başlar ve araştırmacılar fenomenlerin hangi koşullar altında ve hangi vargılarla niçin meydana geldiklerini açıklayabilmeye başlarlar.

Grounded teorisinin üçüncü ve son safhası **seçici kodlamadır** ki bu safha bir merkezi kategoriye veya bir çekirdek değişkeni teşhis etmeyi ve bu merkezi kategoriye sistematik olarak ve mantıksal olarak diğer kategorilerle ilişkilendirmeyi ihtiva eder. Merkezi kategori mevcut kategorilerden evrilebilir veya önceki kodlanmış kategorileri içine alan daha yüksek mertebeli bir kategori olabilir. Merkezi kategoriye ve onun diğer kategorilerle olan ilişkilerini (yani, tentatif teori) geçerlemek için, yeni veri seçici olarak örneklenir. Seçici kodlama analizin gidimini sınırlandırır ve analizi hızlandırır. Aynı zamanda, kodlamacı, yeni veriden ortaya çıkabilecek ve ilgilenilen fenomenle ilişkili olabilecek başka kategoriler için de gözünü açık tutmalıdır (açık kodlama). Bunlar başlangıç teorisinin daha da rafine edilmesine yardımcı olur.

Bundan ötürü, açık, eksensel ve seçici kodlama eş-zamanlı olarak yürütülebilir. Yeni verinin kodlanması ve teorisinin rafine edilmesi **teorik doyuma** ulaşıncaya kadar, yani, ilave veri çekirdek kategorilerde veya ilişkilerde herhangi bir marjinal değişiklik hasıl etmeyinceye kadar devam eder.

“Süreğin karşılaştırma” prosesi kategorilerin, ilişkilerin ve interpretasyonların derinleşen bir anlayışla sürekli yeniden-düzenlenmesini ve şu dört aktivite aşaması arasında bir iteratif etkileşimi ifade eder: (1) her bir kategoriye atanan hadiselerin/metinlerin karşılaştırılması (kategoriye geçerlemek için), (2) kategorilerin ve onların sahip oldukları özelliklerin entegre edilmesi, (3) teorisinin sınırlandırılması (çekirdek kavramlara odaklanmak ve daha az münasebetli kavramları görmezden gelmek) ve (4) teorisinin yazılması (kısa not işleme, senaryolaştırma ve diyagramlaştırma gibi teknikleri kullanma). Bir merkezi kategoriye sahip olmak tüm diğer kategorilerin illaki onun etrafında mükemmelce entegre edilebilecekleri manasına gelmez. Çekirdek kategorinin koşulları/interaksiyonları ve vargıları olarak kilit kategorileri teşhis etmek için, Strauss ve Corbin (1990) mesela senaryolaştırma, kısa not alma ve kavram haritalama gibi çeşitli entegrasyon tekniklerini tavsiye etmişlerdir. **Senaryolaştırmada** (*/storylining*), gözlenen fenomenin bir hikayesinin detaylı bir açıklamasını yapmak için, kategoriler ve ilişkiler kullanılır. Kısa notlar, süstantif kavramlar ve bunların aralarında teorik olarak kodlanmış ilişkiler hakkındaki fikirlerin teorileştirilerek yazılmasıdır. Analiz sırasında geliştirilen fikirlerin izlenmesi ve rafine edilmesi için önemlilerdir. **Kısa not işleme** bu kısa notları, kategoriler arasındaki biçimleri ve ilişkileri keşfetmek için kullanma prosesidir. İkiye iki tablolar, diyagramlar veya figürler veya daha başka görselleştirme vasıtaları istihdam edilebilir. Kavram haritalama kavramların veya kavramlar arasındaki ilişkilerin grafiksel bir temsidir (mesela kutuları ve okları kullanarak). Majör kavramlar tipik olarak bir veya birden çok sayfa veya karton levha üzerine yerleştirilirler ve oklar kullanılmak suretiyle birbirlerine bağlanırlar ve gözlenen veriye en iyi uyacak şekilde yeniden düzenlenirler. Bu işlem garafiksel yazılım paketleri ile de yapılabilir.

Grounded teori üretildikten sonra, içsel tutarlılık ve mantık bakımından rafine edilmelidir. Araştırmacılar merkezi yapının ifade edilen karakteristiklere ve yapılaraya sahip olduğundan emin olamlıdır. Eğer değillerse, veri analizi yinelenebilir. Araştırmacı o zaman tüm kategorilere ait karakteristiklerin ve boyutların değişim gösterdiklerinden emin olmalıdır. Mesela, davranış frekansı böyle bir kategori ise, o zaman veri odaklanılan davranışın hem daha yüksek frekanslı hem de daha düşük frekanslı icracılarına ait kanıt temin etmelidir. Nihayetinde, teori ham veri ile karşılaştırılarak, geçerlenmelidir. Eğer teori gözlenen kanıtla çelişiyor ise, böyle çelişkileri veya açıklanamayan varyasyonları telif etmek için, kodlama prosesi yinelenabilir.

İçerik Analizi

İçerik analizi bir metnin içeriğinin nicel veya nitel bir yaklaşımla sistematik analizidir mesla, kim ne dedi, kime dedi, niçin dedi, ne ölçüde dedi ve etkisi nedir). İçerik analizi tipik olarak şu şekilde yürütülür: Birincisi, analiz edilecek çok sayıda metin (mesela, gazete hikayeleri, finansal raporlar, blog postingleri, online görüşmeler ve benzeri) olduğu zaman, araştırmacı metinler popülasyonu içerisinden analiz edilmek üzere seçilmiş bir metinler kümesini örnekleyerek başlar. Örnekleme rastgele yapılmaz. Ondan ziyade, daha uyarlı içeriğe sahip metinler maksatlıca seçilir. İkincisi, araştırmacı her bir metni ayrı analiz üniteleri olarak ele alınacak segmentlere veya “parçalara (*/chunks*)” bölmek için kurallar teşhis eder ve uygular. Bu prosese *üniteleştirme* adı verilir. Mesela, metindeki varsayımlar, etkiler, sağlayıcılar (*/enablers*) ve bariyerler bu üniteleri oluştururlar. Üçüncüsü, araştırmacı üniteleştirilen her metin segmenti için bir veya birden çok sayıda kavram inşa eder ve uygular. Bu prosese *kodlama* adı verilir. Kodlama yaparken, araştırmacının, metni tasnif ederken bulmaya veya

ortaya çıkarmaya çalıştığı temalara dayanan bir kodlama şeması kullanılır. Dördüncüsü, kodlanmış veri, hangi temaların hangi bağlamlarda en yüksek frekanslarla meydana geldiklerini ve birbirleriyle nasıl ilişkili olduklarını belirlemek için, çoğu zaman hem nicel olarak hem nitel olarak analiz edilir.

İçerik analizinin basit bir tipi **sentiment analizidir** ki bu teknik insanların bir nesneye, bir kişiye veya bir fenomene karşı görüşlerini veya tavırlarını yakalamak için kullanılır. Bir politik aday hakkında bir online foruma gönderilen mesajları okumak ve her mesajı pozitif, negatif veya nötr olarak tasnif etmek böyle bir analiz için bir misaldir. Bu vakada, her mesaj bir analiz ünitesini temsil eder. Bu analiz bir bütün olarak örneğin adaya karşı duruşunun pozitif mi yoksa negatif mi yoksa nötr mü olduğunu teşhis etmeye yardım edecektir. Online görüşmelerin içeriğini benzer bir yoldan incelemek başka bir misaldir. Bu analiz manuel olarak yapılabildiği gibi, çok geniş veri kümeleri (metin kayıtlarının milyonlarcası) için, kodlama prosesini otomatikleştirecek ve insanların sentimentlerinin zamanla nasıl dalgalandığının bir kaydını tutacak doğal dil işleme ve metin analitiklerini haiz bilgisayar yazılımları mevcuttur.

İçerik analizine dair sıkça yapılan bir eleştiri, icra edilen analizin başka araştırmacılar tarafından yinelenmesini sağlayacak sistematik bir prosedürler kümesinin eksikliğidir. Schilling (2006)¹⁹ bu eleştiriye, bir spiral model içerisinde farklı içerik analizi prosedürlerini organize etmek suretiyle cevap vermiştir. Bu model metin interpretasyonunda beş düzey veya safhadan meydana gelir: (1) kaydedilen tapelerin içerik analizi için ham metin verisine veya transkriptlere dönüştürülmesi, (2) ham veriyi yoğunlaştırılmış protokollere dönüştürmek, (3) yoğunlaştırılmış protokolleri önsel bir kategori sistemine dönüştürmek, (4) kodlanmış protokolleri üretmek için, önsel kategori sistemini kullanmak ve (5) ilgilenilen fenomen hakkında interpretasyonlar üretmek için, kodlanmış protokolleri analiz etmek.

İçerik analizi birtakım sınırlamalara da sahiptir. Birincisi, kodlama prosesi metin formunda mevcut olan enformasyon ile kısıtlıdır. Mesela, eğer araştırmacı halkın idam cezası hakkındaki görüşlerini öğrenmeye-çalışmakla ilgileniyorsa, fakat herhangi bir metinsel belge mevcut değilse, o zaman analiz yapılabilmaz. İkincisi, örnekleme yanlılığından kaçınmak için, örnekleme dikkatlice yapılmalıdır. Mesela, eğer popülasyonumuz belli bir konuda yayımlanmış araştırma literatürü ise, o zaman yayımlanmamış araştırmaları veya yayımlanmak üzere olan en son çalışmaları sistematik olarak gözden kaçırmış oluruz.

Hermenötik Analiz

Hermenötik analiz içerik analizinin özel bir tipidir ki burada araştırmacı verilen bir metnin subjektif manasının “interpretasyonunu” onun kendi sosyo-tarihsel bağlamı içerisinde yapmaya çabalar. Kodlama prosesi sırasında metin belgelerinin bağlamını ve manasını göz ardı eden grounded teorisinin veya içerik analizinin aksine, hermenötik analiz nitel veriyi analiz etmek için tamamen interpretatif bir tekniktir. Bu metot varsayar ki yazılı metinler bir yazarın deneyimini bir sosyo-tarihsel bağlam içerisinde anlatır ve bu itibarla interpretasyonu da o bağlam içerisinde yapılmalıdır. Bu sebeple, ele alınan fenomenin, kendi bağlamı içerisinde, daha tam bir anlamını geliştirmek için, araştırmacı metnin tekil interpretasyonu (parça) ve bağlamın holistik bir anlamı (bütün) arasında sürekli olarak gidip-gelir (iterasyon) ki bunu Alman filozofu Martin Heidegger *hermenötik çember* olarak adlandırır. *Hermenötik* kelimesi (tekil) belirli bir interpretasyon metoduna veya sürecine atfedilebilir.

¹⁹ Schilling, J. (2006). “On the Pragmatics of Qualitative Assessment: Designing the Process for Content Analysis,” *European Journal of Psychological Assessment* (22:1), 28-37.

Daha genel olarak, hermenötik interpretasyona ait öğrenme-çalışması ve interpretasyonun teorisi ve pratiğidir. Dinsel öğrenme-çalışmalarından ve dilbilimden türetilen geleneksel hermenötik spezial olarak edebiyat, din ve hukuk alanlarındaki yazılı metinlerin interpretasyonuna matuftur (mesela incil hermenötiği gibi). Yirminci yüzyılda, Heidegger, yalnızca onu gözlemekten daha ziyade, sosyal gerçekliği anlamının daha doğrudan, aracısız ve otantik bir yolunun onu deneyimlemek olduğunu öne sürdü ve interpretasyonun yerine varoluşsal anlama odaklanan *felsefi hermenötiği* önerdi. Heidegger metinlerin sadece okuyucuların yazarın deneyimlerini okumalarına yarayan vasıtalar olmadıklarını fakat aynı zamanda yazarın deneyimlerini yeniden yaşamalarına yarayan vasıtalar olduklarını iddia etmektedir. Mesela Hans-Georg Gadamer gibi Heidegger'in öğrencileri tarafından geliştirilen çağdaş veya modern hermenötik daha sonra sosyal deneyimin iletişimini yapmak için yazılı metinlerin sınırlarını incelemiştir ve devamında interpretif proses için bir çerçeve önermiştir. Bu çerçeve yazılı, sözlü ve sözsüz tüm iletişim formlarını kapsar ve mesela ön-faraziyeler, dil strüktürleri (mesela, gramer, söz-dizimi v.b.) ve semiyotik (sembolizm, mecaz, teşbih ve istihza gibi yazılı işaretlere ait öğrenme-çalışması) gibi yazılı metinlerin iletişimsel kabiliyetlerini kısıtlayan meseleleri keşfetmeye çalışır. Bazen, hermenötik (*/hermenetics*) terimi yerine yalnızca yazılı metinlerin ve spezial olarak dinsel metinlerin interpretasyonuna veya eleştirel açıklamasına izafe edilen tefsir (*/exegesis*) terimi isabetsizce kullanılmaktadır.

Netice

Nitel araştırma metotlarında kodlama proseslerini otomatikleştirmek için, mesela ATLAS.ti.5, Nvivo ve QDA Miner gibi standart yazılım paketleri kullanılabilir. Bu programlar, kullanıcılar tarafından tanımlanabilen kurallar vasıtasıyla, geniş hacimli metin verisini çabukça ve etkince organize edebilir, içerisinde arama yapabilir, sıralayabilir ve işleyebilir. Böyle otomatikleştirilmiş bir analize kılavuzluk etmek üzere, metin verisinin bir örneğinin başlangıçta yapılacak manuel bir incelemesine dayanarak, metin içerisinde aranacak anahtar kelimeleri ve kodları spesifikleştiren bir kodlama şeması yaratılmalıdır. Kodları daha yüksek mertebeli kodlar veya yapılar içerisinde organize etmek için, bu şema hiyerarşik bir şekilde organize edilebilir. Kodlama şeması, farklı metin örnekleri kullanılarak, isabetlilik (*/accuracy*) ve kifayet (*/adequacy*) bakımından geçerlenmelidir. Mamafih, eğer kodlama şeması yanlışsa veya doğru değilse, tüm metin popülasyonunun analizinin sonucu kusurlu ve intepretasyonu-yapılabilmez olabilir. Öte yandan, kompüter programları belli kelimelerin veya sözlerin veya bu kelime veya sözlerin içerisinde kullanıldıkları bağlamın arkasındaki manayı deşifre edemeyebilirler ki bu da geniş ölçekli nitel analizde anlamlı yanlış-interpretasyona yol açabilir.

Bölüm 14

Nicel Analiz: Betimsel İstatistik

Bir araştırma projesinde toplanan sayısal veri istatistiksel araçları kullanmak suretiyle nicel olarak iki safhada analiz edilebilir. **Betimsel analiz** ilgilenilen yapıların veya bu yapılar arasındaki ilişkilerin istatistiksel olarak betimlenmesine, toplaştırılmasına ve sunulmasına matuftur. **Çıkarımsal analiz** hipotezlerin istatistiksel olarak test edilmesine matuftur (teori test etme). Bu bölümde, betimsel analiz için kullanılan istatistiksel teknikleri inceleyeceğiz ve bir sonraki bölümde çıkarımsal analiz için istatistiksel teknikleri inceleyeceğiz. Günümüzde nicel veri analizi çoğunlukla mesela SPSS veya SAS gibi kompüter yazılım paketleri vasıtasıyla icra edilmektedir. Bu bölümde betimlenen kavramları anlayabilmesi için, okuyucuya bu yazılım paketlerinden birisiyle kendisini tanıştırmayı tavsiye edilir.

Veri Hazırlama

Araştırma projelerinde, veri çeşitli kaynaklardan toplanabilir: posta yoklamaları, görüşmeler, ön-test veya son-test deneysel veri, gözlemsel veri ve benzeri. SPSS veya SAS gibi kompüter yazılımları tarafından analiz edilebilmesi için, toplanan verinin mesela bir veri-sayfasında veya başka bir elektronik dosyada makina tarafından okunabilir bir sayısal formata dönüştürülmelidir. Veri hazırlama genellikle aşağıdaki adımları takip eder.

Veri kodlama. Kodlama veriyi sayısal formata çevirme prosesidir. Kodlama prosesine kılavuzluk etmesi için, bir kod-sayfası (*/codebook*) hazırlanmalıdır. Bir **kod-sayfası** bir araştırma öğrenme-çalışmasında yer alan her bir değişkenin, o değişkene ait maddelerin veya ölçülerin (*/items or measures*), her bir maddenin formatının (sayısal, metin, vd.), her bir maddenin yanıt ölçeğinin (yani, adlandırma, sıralama, aralık veya oran ölçeğinde mi ölçülmüş; bu ölçek beşli mi, yedili mi veya başka tip bir ölçek mi?) ve her bir değer sayısal bir formatta nasıl kodlanacağını detaylı betimlemesini içeren komprehensif bir belgedir. Mesela, eğer “kesinlikle katılmıyorum”dan “kesinlikle katılıyorum”a kadar giden çapalara (yani, kategorilere) sahip yedi-puanlı bir ölçekle ölçülen bir maddemiz varsa, bu maddeyi şöyle kodlayabiliriz: “kesinlikle katılmıyorum” için 1, nötr için 4, “kesinlikle katılıyorum” için 7 ve aralardaki çapalar için de sırasıyla 2, 3, 5 ve 6. Mesela sektör tipi gibi adlı veri yine bir kodlama şeması kullanılmak suretiyle, sayısal formda mesela şöyle kodlanabilir: imalat için 1, perakende için 2, finans için 3, sağlık için 4 ve saire. Ancak, elbette ki adlı veri istatistiksel olarak analiz edilebilmez. Mesela yaş, gelir veya test skorları gibi oran ölçeği verisi yanıtlayıcının girdiği üzere kodlanabilir. Bazen, veri toplamak için kullanılan formattan başka farklı bir formda veriyi toplaştırmak ihtiyacı hasıl olabilir. Mesela, “kompüterlerin faydaları” gibi bir yapıyı ölçmek için, eğer bir yoklamada yanıtlayıcılara içerisinden seçim yapabilecekleri faydaların bir işaretleme-listesi (yani, yazılan faydalardan istedikleri kadarını seçebilecekleri) temin edilirse, o zaman işaretlenen maddelerin toplam sayısı toplasık bir fayda ölçüsü olarak kullanılabilir. Şunu ifade etmek te yarar var ki mesela görüşme transkriptleri gibi diğer veri formları istatistiksel analiz için sayısal bir formata çevrilemezler. Kodlama prosesinin farklı kişiler tarafından yürütüldüğü, çok sayıda değişkeni ve ölçüm maddesini ihtiva eden geniş ve kompleks öğrenme-çalışmalarında, kodlama bilhassa önemlidir. Kodlama takımının veriyi tutarlıca kodlaması ve aynı zamanda kodlanan verinin başkalarınca anlaşılması ve interpretasyonunun yapılması temin edilmelidir.

Veri girişi. Kodlanan veri bir veri-sayfasına, veri-tabanına, metin dosyasına veya doğrudan SPSS gibi bir istatistiksel kompüter programına girilebilir. Çoğu istatistik programlarının veri

girişi için kendine ait bir veri editörü mevcuttur. Mamafih, bu programlar veriyi kendilerine ait bir formatta depolarlar (mesela, SPSS veriyi .sav uzantılı dosyalalarda depolar) ki bu da girilen verinin başka istatistik programları ile paylaşılmasını zorlaştırır. Bundan ötürü, veriyi, çoğu zaman, bir veri-sayfasına veya bir veri-tabanına girmek daha iyidir. Böylece, ihtiyaç hasıl olduğunda, veriyi yeniden organize edebiliriz, başka programlara aktarabiliriz ve analiz etmek için verinin alt kümelerini çıkarıp-alabiliriz. Nispeten daha dar veri kümeleri mesela Microsoft Excel gibi bir veri sayfasında depolanabilirken, milyonlarca gözlem ihtiva eden daha geniş veri kümelerini bir veri-tabanına kaydetmek gerekir. Bir veri-sayfasında, her gözlem bir satıra girilebilir ve her ölçüm maddesi bir sütun ile temsil edilebilir. Girilen veri, isabetlik için, giriş esnasında ve sonrasında, maddelere veya gözlemlere gelişigüzel yeniden bakmak suretiyle sık sık kontrol edilmelidir. Dahası, veriyi girerken, kodlayıcı mesela yanıtlayıcının içeriğine bakmaksızın tüm maddeler için “kesinlikle katılıyorum” yanıtını seçmesi gibi kötü veriye açıkça delalet eden durumlar için gözünü açık tutmalıdır. Ters-kodlanmış maddeler de buna dahildir. Böyle bir halde, veri girilmelidir fakat daha sonraki analize dahil edilmemelidir.

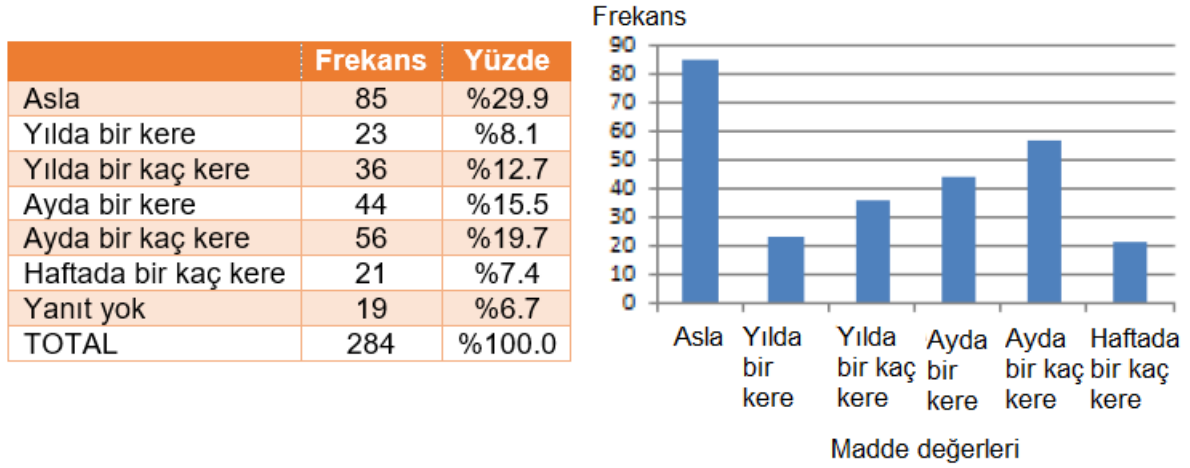
Kayıp değerler. Kayıp veri ampirik bir veri kümesinin kaçınılmaz bir parçasıdır. Eğer lafzı müphem ise veya çok duyarlı ise, yanıtlayıcılar belli sorulara yanıt vermeyebilirler. Böyle problemler plot öğrenme-çalışmalarında erkence belirlenmeli ve esas veri toplama prosesi başlamadan önce düzeltilmelidir. Veri girişi sırasında, bazı istatistik programları boş bırakılan alanları otomatik olarak kayıp değerler olarak görürken, bazıları da kayıp değeri belirtmek için mesela -1 veya 999 gibi spesifik bir sayısal değer girmeyi gerektirir. Veri analizi sırasında, pek çok kompüter programının kayıp değerleri ele almaktaki default modu tek bir kayıp değeri içeriyor olsa bile tüm gözlemi basitçe elemine etmektir. Bu tekniğe *liste-boyu silme* (*/listwise deletion*) adı verilir. Bu şekilde silmek örnek hacmini anlamlıca azaltabilir ve küçük etkileri belirlemeyi aşırı derecede zorlaştırabilir. Bundan ötürü, bazı kompüter programları kayıp değerleri, *impütasyon* adı verilen bir proses vasıtasıyla kestirilmiş bir değer ile ikame etme opsiyonu sunarlar. Mesela, kayıp değer eğer çok maddeli bir ölçekteki tek bir madde ise, izafe edilen değer yanıtlayıcının o ölçeğin geri kalan maddelerine vermiş olduğu yanıtlarının ortalaması olabilir. Eğer kayıp değer rastgele bir mahiyetten daha ziyade sistematik bir mahiyetteyse, böyle bir impütasyon yanlı olabilir. Nispeten daha yansız kestirimler üretebilen iki impütasyon metodu maksimum olabilirlik prosedürleri ve çoklu impütasyon metotlarıdır ki bunların her ikisi de mesela SPSS ve SAS gibi popüler kompüter programları tarafından desteklenmektedir.

Veri taransformasyonu. Bazen, veri değerlerinin manalıca interpretasyonunu yapabilmek için, önce onları trasformasyona tabi tutmamız lazım gelir. Mesela, ters kodlanmış maddeler olabilir ki bunlar ait oldukları yapıyla ters manadadır. Ters kodlanmamış maddelerle karşılaştırmak veya birleştirebilmek için, bunların ters-düz edilmeleri gerekir. Mesela, 1-7 aralığındaki bir ölçekte, gözlenen her bir değeri 8 den çıkarmak o değeri ters-düz edecektir. Transformasyonun başka çeşitleri arasında maddeleri toplayarak yeni maddeler yaratmayı, gözlenen ölçülerin bir kümesinden ağırlıklı bir endeks yaratmayı ve çoklu değerleri daha az sayıda kategoriye indirgemeyi (mesela, gelirleri gelir aralıklarına indirgemek) sayabiliriz.

Tek-değişkeli Analiz

Tek-değişkeli analiz veya başka bir deyişle tek bir değişkenin analizi bir değişkeni genel hususiyetleri bakımından betimlemeye yarayan bir takım istatistiksel tekniklere matuftur. Tek-değişkeli istatistikler şunları kapsar: (1) Frekans dağılımı, (2) merkezileşme ve (3) yayılım. Bir değişkenin **frekans dağılımı** o değişkene ait gözlem değerlerinin veya değer aralıklarının frekanslarının (veya yüzdelilerinin) bir özetidir. Mesela, “dindarlığın” bir ölçüsü olarak, yanıtlayıcıların bir örneğinin dini ritüellere hangi sıklıkta katıldıklarını ölçmek için şu kategorik

ölçeği kullanabiliriz: asla, yılda bir, yılda birkaç, ayda bir, ayda birkaç, haftada bir, haftada birkaç. Eğer her kategoriye düşen gözlemleri sayarsak (veya yüzdesini hesaplırsak) ve bunu Figür 14.1 de gösterildiği üzere bir tablo formunda görselleştirirsek, buna işte frekans dağılımı deriz. Tabloya eklenmiş olan “yanıt yok” kategorisi gerçekten bir kategori olmaktan ziyade, kayıp gözlemlerin sayısını vermektedir. Bu dağılım aynı zamanda Figür 14.1 in sağ tarafında gösterildiği üzere bir çubuk grafik formunda da görselleştirilebilir ki burada yatay eksen değişkenin kategorilerini temsil ederken, dikey eksen her bir kategorinin frekansını veya o kategoriye düşen gözlemlerin yüzdesini temsil eder.



Figür 14.1 Dindarlığın frekans dağılımı

Gözlemlerin bağımsız ve rastgele yapılmış olduğu çok geniş örnekler söz konusu ise, frekans dağılımı, eğer dış kenarlarına bir çizgi çekilirse, çan şekline benzeyen bir görünüme sahip olma eğilimi gösterir. Figür 14.2 de gösterildiği üzere, gözlemlerin çoğu değerlerin gidiminin (*range*) merkezi etrafında öbekleşir ve gidimin aşırı uçlarına doğru gözlemler gittikçe daha azalır. Böyle bir eğriye *normal dağılım eğrisi* adı verilir.

Merkezileşme değerlerin dağılımının merkezine ait bir kestirimdir. Merkezileşmenin başlıca üç kestirimi vardır: aritmetik ortalama (*/mean*), ortanca (*/median*) ve mod (*/mode*). **Aritmetik ortalama** verilen bir dağılıma ait tüm değerlerin basit ortalamasıdır. Şu sekiz test skorundan oluşan veri kümesini ele alalım: 15, 22, 21, 18, 36, 15, 25, 15. Bu değerlerin aritmetik ortalaması $(15 + 22 + 21 + 18 + 36 + 15 + 25 + 15)/8 = 20.875$. Aritmetik ortalamanın *geometrik* (bir dağılımdaki n adet değerlerin çarpımlarının n inci kökü) ve *harmonik* (bir dağılımdaki değerlerin terslerinin aritmetik ortalamasının tersi) tipleri de mevcuttur fakat bu aritmetik ortalamalar sosyal araştırma verisinin istatistiksel analizi için çok popüler değillerdir.

Merkezileşmenin ikinci ölçüsü olan **ortanca** bir dağılımdaki değerlerin gidimi içerisinde ortaya düşen değerdir. Bir dağılımdaki tüm değerleri küçükten büyüğe doğru sıraladıktan sonra, orta sıradaki değeri seçmek suretiyle hesaplanır. Eğer orta sırada iki değer varsa (yani bir dağılımdaki değerlerin sayısı bir çift sayı ise), ortada kalan iki değer aritmetik ortalaması ortancayı temsil eder. Yukarıdaki misalde, sıralanmış değerler şunlardır: 15, 15, 15, 18, 21, 22, 25, 36. Ortaya düşen iki değer 18 ve 22 olduğu için, ortanca $(18 + 22)/2 = 20$.

Son olarak, **mod** bir dağılımdaki değerler içerisinde en yüksek frekansla gözlenen değerdir. Bir önceki misalde, en yüksek frekansla gözlenen değer 15 tir, öyle ki bu değer yukarıdaki test skorları kümesinin modudur. Aklından çıkarmamalısın ki mesela aritmetik ortalama, ortanca, mod ve benzerleri gibi bir örnekten kestirilen herhangi bir değer bir **istatistik** olarak adlandırılır.

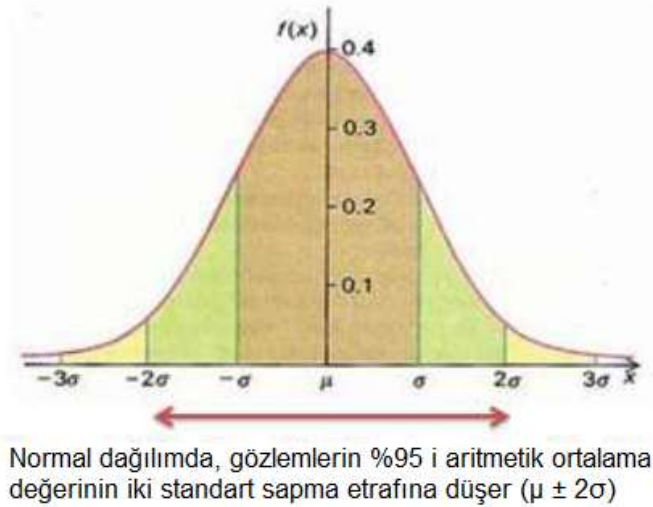
Yayılmış değerlerin merkezileşmenin etrafında ne şekilde saçıldığına matuftur. Mesela, değerler

aritmetik ortalamasının etrafında ne kadar sıkıca veya ne kadar gevşekçe öbikleşiyor. Yayılımın müştereken kullanılan iki ölçüsü gidim (*range*) ve standart sapmadır. **Gidim** bir dağılımdaki değerlerin en büyüğü ve en küçüğü arasındaki farktır. Bir önceki misalimizin gidimi $36 - 15 = 21$.

Gidim bilhassa dışlakların (*outliers*) varlığına karşı duyarlıdır. Mesela, eğer yukarıdaki dağılımda en büyük değer 85 olsaydı ve diğer değerler aynı kalsaydı, gidim $85 - 15 = 70$ olurdu. İkinci yayılım ölçüsü olan **standart sapma** her değerın dağılımın aritmetik ortalamasına ne kadar yakın veya ne kadar uzak olduğunu hesaba katan bir formülü kullanmak suretiyle dışlakların bu etkisini düzeltir:

$$\sigma = \sqrt{\frac{\sum_{i=1}^N (y_i - \mu)^2}{N}}$$

ki burada σ standart sapma, y_i i inci gözlem (veya değer), μ aritmetik ortalama, N gözlemlerin toplam sayısı ve Σ tüm gözlemler için toplama yapılacağını gösteren matematiksel simgedir. Standart sapmanın karesi bir dağılımın **varyansı** diye adlandırılır. Normal dağılıma sahip bir frekans dağılımında, Figür 14.2 de gösterildiği üzere, gözlemlerin %68 inin aritmetik ortalamasının bir standart sapma civarında ($\mu \pm 1\sigma$) yer aldığı, gözlemlerin %95 inin iki standart sapma civarında ($\mu \pm 2\sigma$) yer aldığı ve gözlemlerin %99.7 sinin üç standart sapma civarında ($\mu \pm 3\sigma$) yer aldığı görülür.



Figür 14.2 Normal dağılım

İki-değişkeli Analiz

İki-değişkeli analiz iki değişkenin birbirleriyle nasıl ilişkili olduğunu inceler. En müştereken kullanılan iki-değişkeli istatistik **iki-değişkeni korelasyondur** ki çoğu zaman kısaca “korelasyon” diye adlandırılır. Korelasyon -1 ve +1 arasında bir rakamdır ve iki değişken arasındaki ilişkinin gücünü belirtir. Diyelim ki 20 yanıtlayıcıdan oluşan bir örnekte yaşınn özgüven ile nasıl ilişkili olduğunu öğrenmeye-çalışmayı arzu ediyoruz. Yani, yaş arttıkça özgüven artıyor mu, azalıyor mu veya değişmeden kalıyor mu? Eğer özgüven artıyorsa, iki değişken arasında pozitif bir ilişki var demektir. Eğer özgüven azalıyorsa, negatif bir korelasyon var demektir. Eğer aynı kalıyorsa, korelasyon sıfırdır. Bu korelasyonun değerini hesaplamak için, Tablo 14.1 deki hipotetik veri kümesini ele alalım.

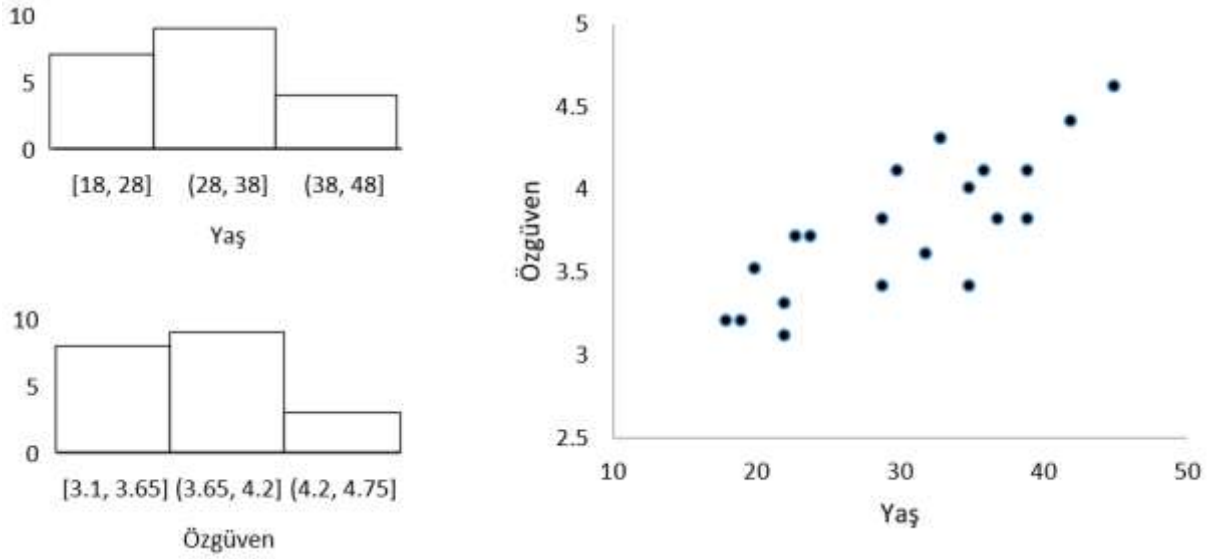
Gözlem	Yaş (x)	Özgüven (y)	xy	x ²	y ²
1	39	4.1	159.9	1521	16.81
2	45	4.6	207	2025	21.16
3	29	3.8	110.2	841	14.44
4	42	4.4	184.8	1764	19.36
5	19	3.2	60.8	361	10.24
6	22	3.1	68.2	484	9.61
7	39	3.8	148.2	1521	14.44
8	30	4.1	123	900	16.81
9	33	4.3	141.9	1089	18.49
10	23	3.7	85.1	529	13.69
11	20	3.5	70	400	12.25
12	18	3.2	57.6	324	10.24
13	24	3.7	88.8	576	13.69
14	22	3.3	72.6	484	10.89
15	29	3.4	98.6	841	11.56
16	35	4.0	140	1225	16.00
17	36	4.1	147.6	1296	16.81
18	37	3.8	140.6	1369	14.44
19	35	3.4	119	1225	11.56
20	32	3.6	115.2	1024	12.96
Topl. (Σ)	609	75.1	2339.1	19799.0	285.45

Tablo 14.1 Yaş ve özgüven değişkenlerine ait hipotetik veri

Bu veri kümesindeki iki değişken yaş (x) ve özgüvendir (y). Yaş bir oran-ölçeği değişkeniyken, özgüven bir ortalama skordur öyle ki “kesinlikle katılmıyorum”dan “kesinlikle katılıyorum”a giden bir 7-noktalı Likert ölçeği kullanılarak ölçülen çok-maddeli bir özgüven ölçeğinden hesaplanır. Her bir değişkene ait histogramlar Figür 14.3 ün sol tarafında gösterilmektedir. Bu histogramlara ayrı ayrı baktığımızda, iki değişken arasında potansiyel bir ilişkinin varlığına veya yokluğuna dair herhangi bir belirti göstermiyorlar. İki-değişekli korelasyonu hesaplamak için formül şudur:

$$r_{xy} = \frac{\sum x_i y_i - n \bar{x} \bar{y}}{(n-1) s_x s_y} = \frac{n \sum x_i y_i - \sum x_i \sum y_i}{\sqrt{n \sum x_i^2 - (\sum x_i)^2} \sqrt{n \sum y_i^2 - (\sum y_i)^2}}$$

ki burada r_{xy} korelasyon, $\bar{x}$ ve $\bar{y}$ iki değişkenin örnek aritmetik ortalamaları ve s_x ve s_y standart sapmalarıdır. Yukarıdaki formül kullanılmak suretiyle, Tablo 14.1 de gösretildiği üzere, yaş ve özgüven arasındaki korelasyon elde 0.79 olarak hesaplanır. Bu rakam şunu belirtir ki yaşın özgüven ile güçlü pozitif bir korelasyonu vardır. Yani, yaş arttıkça özgüven de artma eğilimi göstermektedir ve yaş azaldıkça da azalma eğilimi. Bu biçimi Figür 14.3 ün sağ tarafındaki saçılım grafiğinde görmekteyiz. Bu iki-değişekli saçılım grafiği esasen yatay eksendeki yaşa karşı dikey eksendeki özgüvenin bir grafiğidir. Bu grafik kabaca sol alt köşeden sağ üst köşeye doğru giden eğimli bir düz çizgiye benzemektedir (yani, pozitif eğimli), öyle ki pozitif bir korelasyonu belirtmektedir. Eğer iki değişken negatif korelasyonlu olsalardı, saçılım grafiği sol üst köşeden sağ alt köşeye doğru eğimli olurdu (yani, negatif eğimli) ve yaştaki bir artışın özgüvende bir azalmaya tekabül edeceği manasına gelirdi. Eğer iki değişken korelasyonsuz olsalardı, saçılım grafiği yaklaşık yatay bir düz çizgiye benzerdi (yani, sıfır eğimli) ve yaştaki bir artışın özgüvende herhangi bir sistematik karşılığı olmadığı manasına gelirdi.



Figür 14.3 Yaşın ve özgüvenin histogramları (solda) ve saçılım grafiği (sağda)

İki-değişkeli korelasyonu hesapladıktan sonra, araştırmacılar çoğu zaman bu korelasyonun anlamlı bir korelasyon mu (yani, hakikaten var olan) yoksa sadece şans eseri mi olduğunu bilmek isterler. Korelasyonun anlamlı olması demek örnekten hesaplanan r korelasyonunun bu örneğin içinden çekilmiş olduğu popülasyondaki hakiki ρ korelasyonunun isabetli bir kestirimi olması demektir. Bu soruyu cevaplamak şu hipotezlerin testini gerektirir:

$$H_0: \rho = 0$$

$$H_1: \rho \neq 0$$

H_0 *sıfır hipotezi* olarak adlandırılır ve H_1 *alternatif hipotez* olarak adlandırılır (bazen H_A olarak ta yazılabilir). Her ne kadar iki hipotez olarak görünseler bile, esasen H_0 ve H_1 tek bir hipotezi temsil ederler çünkü bunlar birbirinin tam karıştıdır. Biz H_0 dan ziyade H_1 i test etmek isteriz. Yine dikkat etmelisin ki H_1 yönsüz bir hipotez çünkü r nin sıfırdan büyük mü yoksa küçük mü olduğunu spesifikleştirmiyor. Yönlü hipotezler $H_0: \rho \leq 0$; $H_1: \rho > 0$ (eğer pozitif bir korelasyon için test yapıyorsak) olarak spesifikleştirilir. Yönlü bir hipotez için anlamlılık testi tek-yönlü t -testi kullanılarak yapılırken, yönsüz bir hipotez için anlamlılık testi çift-yönlü t -testi kullanılarak yapılır.

İstatistiksel testlerde, alternatif hipotez doğrudan test edilebilmez. Onun yerine, sıfır hipotezini belli bir olasılık düzeyinde reddetmek suretiyle, dolaylı olarak test edilir. İstatistiksel testler her zaman olasılıksaldır çünkü örnek veriye dayanan çıkarımlarımızın popülasyon için de geçerli olduğundan asla emin olamayız çünkü örneğimiz asla popülasyona eşit değildir. İstatistiksel bir çıkarımın salt şans eseri olması olasılığı **P -değeri** olarak adlandırılır. P -değeri bir **anlamlılık düzeyi** (α) ile karşılaştırılır. Anlamlılık düzeyi (α) çıkarımımızın yanlış olması halinde göze aldığımız maksimum risk düzeyini temsil eder. İstatistiksel analizlerin pek çoğunda, α nın değeri 0.05 olarak tespit edilir. Eğer bir P -değeri $\alpha = 0.05$ ten daha küçük ise, sıfır hipotezini reddetmek için ve dolayısıyla alternatif hipotezi kabul etmek için yeterli istatistiksel kanıtımız olduğunu belirtir. Eğer $P > 0.05$ ise, o zaman sıfır hipotezini reddetmek için ve dolayısıyla alternatif hipotezi kabul etmek için yeterli istatistiksel kanıtımız yoktur.

Yukarıda hipotezi test etmenin en kolay yolu hesapladığımız r değerini arzulanan anlamlılık düzeyi (α), serbestlik derecesi (df) ve arzulanan testin tek-yönlü mü yoksa çift-yönlü mü olacağına bağlı olarak oluşturulan tablolarda yer alan kritik r değeri ile karşılaştırmaktır. Bu tabloları istatistik kitaplarında veya internette bulabiliriz. (Aynı zamanda, anlamlılık testi için

bir kompüter yazılımı kullanabiliriz.) **Serbestlik derecesi** bir istatistiğin hesaplanmasında bağımsızca değişebilen gözlem değerlerinin sayısıdır. Korelasyon için, df basitçe $n - 2$ ye eşittir; yani, Tablo 14.2 deki veri için $20 - 2 = 18$. Tek-yönlü ve çift-yönlü testler için ayrı ayrı istatistiksel tablolar oluşturulmuştur. Çift-yönlü tabloda, $\alpha = 0.05$ ve $df = 18$ durumunda r için kritik değer 0.44 tür. Hesaplanan bir korelasyonun anlamlı olması için, bu 0.44 lük kritik değerden daha büyük olması veya -0.44 ten daha küçük olması gerekir. Bizim hesapladığımız 0.79 luk değer 0.44 ten daha büyük olduğu için, yaş ve özgüven arasında anlamlı bir korelasyonun mevcut olduğu neticesine varırız. Yani, kendi veri kümemizden hesapladığımız bu korelasyonun şans eseri olması olasılığı %5 ten daha azdır. Bundan ötürü, $\rho = 0$ olduğunu iddia eden sıfır hipotezini reddedebiliriz ki bu da dolaylı olarak $\rho \neq 0$ olduğunu iddia eden alternatif hipotezin olasılıkla doğru olduğu manasına gelir.

Araştırma öğrenme-çalışmalarının çoğu ikiden daha çok sayıda değişkeni ihtiva eder. Eğer n sayıda değişken varsa, bu n değişken arasında $n(n - 1)/2$ sayıda olanaklı korelasyona sahip olabiliriz. Böyle korelasyonları korelasyon formülünü kullanarak elle hesaplamak (Tablo 14.1 de yaptığımız gibi) yerine, mesela SPSS gibi kompüter yazılımlarını kullanarak kolayca hesaplayabiliriz. Bu yazılımlar hesaplanan korelasyonları Tablo 14.2 de gösterildiği gibi bir korelasyon matrisi formunda bize verirler. Bir korelasyon matrisi birinci satırı ve birinci sütunu boyunca değişken adlarının listelendiği ve değişken çiftleri arasındaki korelasyonların da bu değişkenlerin adlarının çakıştığı hücelere yazıldığı bir matristir. Bu matrisin baş diyagonalinde (sağ üst köşeden sol alt köşeye) yer alan değerler daima 1 dir çünkü her değişkenin kendisiyle korelasyonu mükemmeldir. Dahası, korelasyon yönsüz olduğu için, V1 ve V2 değişkenleri arasındaki korelasyon V2 ve V1 arasındaki korelasyonla aynıdır. Bundan ötürü, alt üçgen matris (baş diyagonalin altındaki değerler) üst üçgen matrisin (baş diyagonalin üstündeki değerler) bir aynadan yansımasıdır ve bundan ötürü çoğu zaman basitleştirmek için yalnızca alt üçgen matrisi yazarız. Eğer korelasyonlar aralık ölçeğinde ölçülmüş değişkenlere aitse, o zaman bu spesifik tipteki korelasyonlara **Pearson çarpım moment korelasyonu** adı verilir.

	V1	V2	V3	V4	V5	V6	V7	V8
V1	1.000							
V2	0.274	1.000						
V3	-0.134	-0.269	1.000					
V4	0.201	-0.153	0.075	1.000				
V5	-0.095	-0.166	0.278	-0.011	1.000			
V6	-0.129	0.280	-0.348	-0.378	-0.009	1.000		
V7	0.171	-0.122	0.296	0.086	0.193	0.233	1.000	
V8	0.518	0.238	0.238	-0.227	-0.551	0.082	-0.102	1.000

Tablo 14.2 Sekiz değişken için hipotetik bir korelasyon matrisi

İki-değişkeli veriyi sunmanın başka bir yararlı yolu da çapraz-tablolarıdır (bazen daha formal olarak kontenjans tablosu denir). Bir **çapraz-tablo** öyle bir tablodur ki iki veya daha çok sayıda nitel değişkenin kategorilerinin tüm kombinasyonlarının frekanslarını (veya yüzdelerini) betimler. Nitel değişkenler adlandırma ölçeğinde ve/veya sıralama ölçeğinde ölçülmüş olabilirler. Bir misal olarak, farzedelim ki 20 öğrenciye ait Tablo 14.3 teki cinsiyet ve harf-notu gözlemlerine sahibiz. Cinsiyet bir adlandırma değişkenidir (erkek/dişi veya E/D) ve harf-notu üç düzeyi (A, B ve C) olan bir sıralama değişkenidir. Verinin basit bir çapraz-tablosu, 2×3 bir matriste, cinsiyetin ve harf-notunun ortak dağılımını gösterecektir (yani, mutlak frekanslar cinsinden veya yüzdeler cinsinden, her harf-notu kategorisinde her cinsiyet kategorisinden kaç öğrenci olduğunu). Bu matris bizim A, B ve C harf-notlarının erkek ve dişi öğrenciler arasında

eşit dağılıp dağılmadığını görmemize yardımcı olacaktır. Tablo 14.3 teki çapraz-tablo verisi A harf-notlarının dağılımının dişi öğrenciler lehine ağırca yanlı olduğunu gösteriyor: 10 erkek ve 10 dişi öğrenciden müteşekkil bir örnekte yalnızca 1 erkek öğrencinin A harf-notu almasına karşılık 5 dişi öğrenci A harf-notu almıştır. Tersine, C harf-notlarının dağılımı erkek öğrenciler lehine yanlıdır: yalnızca 1 dişi öğrencinin C harf-notu almasına karşılık 3 erkek öğrenci C harf-notu almıştır. Mamafih, B harf-notlarının dağılımı bir şekilde ünitordur öyle ki 6 erkek öğrenci ve 5 dişi öğrenci. Bu tablonun son satırı ve son sütunu marjinal toplamalar (veya kenar toplamaları) olarak adlandırılır çünkü her bir kategoriye ait toplamaları belirtirler ve tablonun kenarlarında gösterilirler.

Gözlem	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Cinsiyet	D	E	D	E	D	E	E	E	D	D	E	E	E	D	D	E	D	D	D	E
H. notu	A	B	B	B	C	A	C	B	B	A	B	C	C	B	B	B	A	A	B	B

Hipotetik veri kümesi

			Harf notları			Toplam
			A	B	C	
Cinsiyet	Erkek	Sayı	1	6	3	10
		Beklenen sayı	2.5	5.5	2.0	
	Dişi	Sayı	4	5	1	10
		Beklenen sayı	2.5	5.5	2.0	
Toplam			5	11	4	20

Yaşa karşı cinsiyetin çapraz-tablosu

Tablo 14.3 Misal çapraz-tablo analizi

Her ne kadar Tablo 14.3 te erkek ve dişi öğrenciler arasında harf-notu dağılımına ait belirgin bir biçim görebiliyor olsak ta, bu hakiki mi veya “istatistiksel olarak anlamlı mı”? Bir başka deyişle, yukarıdaki frekanslar arasındaki farklar salt şansa bağlı olarak gerçekleşmesi beklenenden farklı mı? Bu soruyu cevaplamak için, 2×3 çapraz-tablo matrisinin her bir hücresindeki gözlemlerin beklenen sayısını hesaplamalıyız. Bu, her bir hücre için marjinal satır toplamını marjinal sütun toplamı ile çarpıp, gözlemlerin toplam sayısına bölmek suretiyle yapılır. Mesela, erkek ve A harf-notu hücresi için beklenen sayı $= \frac{5(10)}{20} = 2.5$. Başka bir deyişle, A harf-notu alacak erkek öğrencilerin sayısını 2.5 olarak beklerdik fakat gerçekte yalnızca 1 öğrenci A harf-notu almıştır. Beklenen ve gözlenen sayılar arasındaki bu farkın anlamlı olup olmadığı bir *ki-kare testi* kullanılmak suretiyle test edilebilir. Ki-kare istatistiği tüm hücrelerdeki gözlenen ve beklenen sayılar arasındaki farkların aritmetik ortalaması olarak hesaplanabilir. Ondan sonra, bu sayıyı arzu edilen anlamlılık düzeyine (α) ve serbestlik derecesine (df) tekabül eden kritik değer ile karşılaştırabiliriz. Bir çapraz-tablo için serbestlik decesi basitçe $(m - 1)(n - 1)$ olarak hesaplanır ki burada m ve n sırasıyla satır ve sütun sayılarıdır. Bu misalde, $df = (2 - 1)(3 - 1) = 2$. Herhangi bir istatistik kitabında bulabileceğimiz standart ki-kare tablosuna baktığımızda, $p = 0.05$ ve $df = 2$ için kritik k-kare değeri 5.99 dur. Bizim gözlenen verimizden hesaplanan ki-kare değeri 1.00 dır öyleki bu değer kritik değerden daha küçüktür. Bundan ötürü, gözlenen harf-notu dağılımının salt şans eseri olması beklenenden dağılımdan istatistiksel olarak farklı olmadığı neticesine varmalıyız.

Bölüm 15

Nicel Analiz: Çıkarımsal İstatistik

Çıkarımsal istatistikler değişkenler arasındaki çağrışımlar hakkında neticeler çıkarmak için kullanılan istatistiksel prosedürlerdir. Şu bakımdan betimsel istatistiklerden ayrılırlar ki hipotezleri test etmek için münhasıran dizayn edilmişlerdir. Bu kategoride telakki edilecek pek çok istatistiksel prosedür mevcuttur ki bunların pek çoğu mesela SPSS ve SAS gibi modern istatistiksel yazılımlar tarafından desteklenir. Bu bölüm sadece en temel ve yaygınca kullanılan prosedürlere kısa bir giriş mahiyetindedir; daha ileri prosedürler için okuyucunun formal istatistik kitaplarına başvurması tavsiye edilir.

Temel Kavramlar

İngiliz felsefeci Karl Popper (1902-1994) teorilerin asla ispatlanamayacağını, yalnızca yanlışlanabileceğini söylemiştir. Bir misal olarak, yarın güneşin doğacağını nasıl ispatlayabiliriz? Popper demiştir ki hatırladığınız her gün güneşin doğmuş olması illaki yarın da doğacağı manasına gelmez. Çünkü, tümevarımsal olarak türetilen teoriler sadece bir takım konjektürlerdir ki müstakbel fenomenin tahminçileri olabilirler veya olmayabilirler. Bunun yerine, Popper şunu önermiştir: illaki ispatlamaksızın, güneşin hergün doğacağı teorisini varsayalım ve eğer belli bir gün güneş doğmaz ise, bu teoriyi yanlışlayalım ve reddedelim. Benzerce, hipotezleri karşı kanıtı dayanarak ancak reddedebiliriz fakat onları asla hakikaten kabul edemeyiz çünkü kanıtın mevcudiyeti daha sonra karşı kanıtın gözlenebilmeyeceği manasına gelmez. İlgilendiğimiz bir hipotezi hakikaten kabul edebilmeyeceğimiz için (alternatif hipotez), alternatif hipotezin karşıtı olarak bir sıfır hipotezi formüle ederiz ve ondan sonra sıfır hipotezini reddetmek için ampirik kanıt kullanırız. Böylece alternatif hipotezimiz için dolaylı olasılıksal destek sağlamış oluruz.

Sosyal bilim araştırmasında hipotezleştirilen ilişkileri test etmeye dair ikinci problem şudur ki yanıt değişkeni (veya bağımlı değişken) sonsuz sayıda harici değişkenin tesiri altında olabilir ve tüm bu harici etkileri ölçmek ve kontrol altına almak akıl karı değildir. Bu sebeple, gözlenen bir örnekte iki değişken ilişkili görünebilseler dahi, popülasyonda hakikaten ilişkili olmayabilirler ve bu yüzden çıkarımsal istatistikler asla kesin veya belirlemeci değildir, fakat her zaman olasılıksallardır.

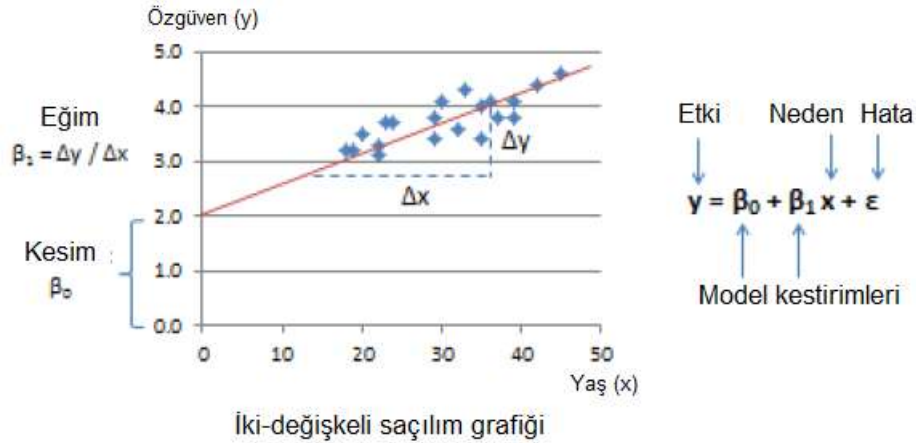
Gözlenen bir örnekte iki değişken arasındaki bir ilişkinin anlamlı olduğunu ve şans eseri olmadığını nasıl biliriz? Sir Ronald A. Fisher (1890-1962), tarihte en çok öne çıkan istatistikçilerden biri, anlamlılık testi için temel taşlarını döşemiştir. Fisher demiştir ki eğer bir istatistiksel sonucun şans eseri reddedilme olasılığının %5 veya daha az olduğu gösterilebilirse, o istatistiksel sonuç anlamlı addedilebilir. Çıkarımsal istatistikte, bu olasılık **P-değeri** olarak adlandırılır, %5 anlamlılık düzeyi olarak adlandırılır (α) ve P -değeri ve α arasında arzulanan ilişki $P \leq 0.05$ olarak gösterilir. **Anlamlılık düzeyi** örnekten popülasyona yaptığımız çıkarımın bedeli olarak göze aldığımız maksimum risk düzeyidir. Eğer P -değeri 0.05 veya %5 ten daha küçükse, bunun manası şudur: Sıfır hipotezini yanlışlıkla reddetme şansımız %5 tir ki buna Tip I hata denir. Eğer $P > 0.05$ ise, sıfır hipotezini reddetmek için veya alternatif hipotezi kabul etmek için yeterli kanıtımız yoktur.

Aynı zamanda, birbiriyle ilişkili üç istatistiksel kavramı da anlamalıyız: örnekleme dağılımı, standart hata ve güven aralığı. Bir örnekleme dağılımı öğrenme-çalışmamızda ilgilendiğimiz popülasyondan çekilen sonsuz sayıdaki aynı hacimli örnekten hesaplanan bir istatistiğin teorik

dağılımıdır. Mamafih, bir örnek hiç bir zaman bir popülasyonun özdeşi olmadığı için, her örnek **standart hata** adı verilen bir hata düzeyini daima özünde barındırır. Eğer bu standart hata küçükse, o zaman örnekten türetilen istatistiksel kestirimler (mesela örnek aritmetik ortalaması gibi) popülasyonun makul kestirimleridir. Örnek kestirimlerimizin kesinliği **güven aralığı** (CI) cinsinden tanımlanır. Bir %95 lik güven aralığı, örnek kestirime iki standart hata ekleyerek ve çıkararak elde edilen aralık olarak tanımlanır. Buna göre, gözlediğimiz örneğe ait kestirimin %95 lik güven aralığından söz ettiğimizde, bunun manası popülasyon parametresinin gözlediğimiz örneğe ait kestirimin iki standart hata civarında olduğuna %95 güven duyduğumuzdur. *P*-değeri ve güven aralığı beraberce elde ettiğimiz sonucun olasılığı hakkında ve bu sonucun müteakıl popülasyon parametresine yakınlığı hakkında bize çok güzel fikir verir.

Genel Doğrusal Model

Sosyal bilim araştırmasında kullanılan çoğu çıkarımsal istatistiksel prosedürler **genel doğrusal model** (GLM) adı verilen genel bir istatistiksel model ailesinden türetilir. Bir *model* bir veri kümesini temsil etmek için kullanılabilen kestirilmiş bir matematiksel eşitliktir. *Linear* ise bir doğruya (yani bir düz çizgiye) atıfta bulunur. Buna göre, bir GLM bir eşitlikler sistemidir öyle ki gözlenen verideki ilişkilerin biçimlerini temsil etmek için kullanılır.



Figür 15.1 Basit doğrusal model

En basit tipteki GLM iki-değişkeli bir doğrusal modeldir öyle ki bir açıklayıcı değişken (veya bağımsız değişken, neden değişkeni, tahminci değişken) ve bir yanıt değişkeni (veya bağımlı değişken, etki değişkeni, çıkım değişkeni) arasındaki ilişkiyi inceler. Yine farzedelim ki bu iki değişken sırasıyla yaş ve özgüven olsun. Bu ilişkinin iki-değişkeli saçılım grafiği Figür 15.1 de gösteriliyor. Yaş (tahminci) yatay ekseninde (veya *x*-ekseninde) ve özgüven (çıkım) dikey ekseninde (veya *y*-ekseninde). Saçılım grafiğine baktığımızda, yaş ve özgüven kombinasyonlarını temsil eden bireysel gözlemlerin genellikle sağ alt köşeden sol üst köşeye doğru giden hayali bir düz çizgi etrafında saçıldıkları görülüyor. Bu düz çizginin (yani doğrunun) parametrelerini yani eğimini ve kesimini GLM ile kestirebiliriz. Lise cebiri kullanarak, biliyoruz ki herhangi bir doğruyu $y = mx + c$ eşitliği ile temsil edebiliriz ki burada *m* doğrunun eğimi (*x* in 1 birim değişmesine karşılık *y* deki değişme) ve *c* kesimidir (*x* sıfırken, *y* nin değeri). GLM de, bu eşitlik formal olarak şöyle temsil edilir:

$$y = \beta_0 + \beta_1 x + \varepsilon$$

ki burada β_0 eğim, β_1 kesim ve ε hata terimidir. ε gerçek gözlemlerin kendi tahmin edilen değerlerinden sapmalarını temsil eder çünkü gözlemlerin pek çoğu yakınında olmakla beraber

tam olarak doğrunun üzerine düşmezler (yani GLM mükemmel değil). Şunu da söyleyelim ki bir doğrusal model ikiden çok sayıda tahminci değişkene sahip olabilir. İki tahminci bir doğrusal modeli görselleştirmek için, üç-boyutlu öyle bir küp hayal et ki bu küpün dikey ekseninde çıkım (y) ve iki yatay ekseninde iki tahminci (diyelim, x_1 ve x_2) yer alsın. İki veya daha çok sayıda değişken arasındaki ilişkiyi betimleyen doğruya regresyon doğrusu denir, β_0 ve β_1 (ve diğer beta katsayıları) *regresyon katsayıları* olarak adlandırılır ve regresyon katsayılarını kestirme prosesine **regresyon analizi** denir. Tahminci sayısı n olan regresyon analizi için GLM şudur:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \dots + \beta_n x_n + \varepsilon$$

Yukarıdaki eşitlikte, x_i tahminci değişkenlerinin (veya yanıt değişkenlerini veya bağımsız değişkenlere) *kovaryete* adı da verilir. Kovaryetelerin bazıları teorik olarak araştırmacının ilgisine mazhar olmayabilirler fakat y yanıt değişkeni üzerinde yine de bir etkileri olabilir. Bu sebeple, kovaryeteler kontrol altına alınmalılardır ki ilgilenilen açıklayıcı değişkenlerin geriye kalan etkisi daha kesin olarak belirlenebilsin. Bir regresyon eşitliğinde kovaryeteler sistematik hatalardan mesulken, hata terimi (ε) rastgele hatalardan mesuldür. GLM de çoğu değişken aralık veya oran ölçeğinde olasılar bile, bu zorunlu değildir. Hatta bazı tahminci değişkenler adlandırma ölçeğinde ölçülen değişkenler dahi olabilir (mesela, cinsiyet: erkek veya dişi) ki bunlar *kukla değişkenler* olarak sayısal kodlanırlar. Kukla değişkenler yalnızca iki olanaklı değer alabilirler: 0 veya 1 (cinsiyet misalinde, “erkek” 0 ile temsil edilebilir ve “dişi” 1 ile veya tersi). Adlandırma ölçeğinde ölçülen ve n kategorisi olan bir değişken $n - 1$ kukla değişken kullanılarak temsil edilebilir. Mesela, endüstri sektörü adlı değişkenin tarım, imalat ve hizmet olmak üzere 3 kategorisi varsa, iki kukla değişkenin (x_1 , x_2) kombinasyonu ile şöyle temsil edilebilir: tarım için (0, 0), imalat için (0, 1) ve hizmet için (1, 1). Kukla değişkenler burada gözlem değerlerinin nitel değişkenin kategorilerine aidiyetleri bakımından bir indikatör vazifesi ifa edecektir, modele nicel bir katkı sağlamaktan daha ziyade (mesela bir deneysel dizayndaki muamele grubu ve kontrol grubu gibi). Böylece, her grubun istatistiksel parametreleri ayrı ayrı kestirilebilecektir.

GLM çok güçlü bir istatistiksel araçtır çünkü yalnızca tek bir istatistiksel metot değildir. GLM çeşitli tiplerdeki tahminci değişkenler ve çıkım değişkenleri ile sofistike analizler yapmak için kullanılabilecek bir metotlar ailesidir. Eğer, kukla değişkenler vasıtasıyla, bir tahminci değişkenin farklı düzeylerinin bir çıkım değişkeni üzerindeki etkilerini karşılaştırıyorsak bir *varyans analizi* (ANOVA) yapıyoruz demektir. Eğer, bir veya birden çok sayıda kovaryetenin etkisini kontrol altına almak suretiyle, bir ANOVA yapıyorsak, bir *kovaryans analizi* (ANCOVA) yapıyoruz demektir. Aynı zamanda, birden fazla yanıt değişkenimiz (y_1 , y_2 , y_3 , ..., y_n) de olabilir ve her bir yanıt değişkeni için farklı bir eşitlikten meydana gelen (her bir eşitliğin kendine mahsus regresyon katsayıları kümesine sahip olduğu) bir “eşitlikler sistemi” kullanarak bunları temsil edebiliriz. Öte taraftan, eğer çıkım değişkenlerini tahminci değişkenlerin aynı kümesiyle tahminlenecek şekilde modellersek, bu analize *çok-değişkeli regresyon* adı verilir. Eğer çoklu çıkım değişkenleri ile ANOVA veya ANCOVA yapıyorsak, bunlara sırasıyla *çok-değişkeli ANOVA* (MANOVA) veya *çok-değişkeli ANCOVA* (MANCOVA) adı verilir. Eğer birbiriyle bağlantılı bir regresyon eşitlikleri sisteminde bir regresyon eşitliğindeki bir çıkım değişkenini diğer bir eşitlikte tahminci değişken olarak modelliyorsak, o zaman *strüktürel eşitlik modellemesi* adı verilen çok sofistike bir analiz tipine sahibiz demektir. GLM de en önemli problem *model spesifikasyonudur*. Yani, ilgilenilen fenomeni en iyi temsil edecek regresyon eşitliğini (veya bir eşitlik sistemini) nasıl spesifikleştireceğimiz. Model spesifikasyonu, gözlenen veriye neyin en iyi uyacağından daha ziyade, öğrenilmeye-çalışılan fenomen hakkındaki teorik mülâhazalara dayanılarak yapılmalıdır. Veri modelin geçerleştirilmesinde rol oynar yoksa modelin spesifikasyonunda değil.

İki-grup Karşılaştırması

En basit çıkarımsal analizlerden biri bir rastgeleleştirilmiş salt son-test kontrol grubu dizaynında muamele grubuna ve kontrol grubuna ait son-test çıkımlarını karşılaştırmaktır. Mesela spesiyal bir matematik programına kaydolan öğrencilerin geleneksel matematik müfredatına kaydolanlara göre daha yüksek performans gösterip göstermedikleri gibi. Bu durumda, tahminci değişken bir kukla değişkendir (muamele grubu için 1, kontrol grubu için 0) ve çıkım değişkeni olan performans oran ölçeğindedir (mesela, spesiyal programı takibeden matematik testinin skoru). Bu basit dizayn için analitik teknik *tek-yönlü ANOVA* (yalnızca bir tahminci değişkeni ihtiva ettiği için tek-yönlü) dir ve kullanılacak istatistiksel test *Student t-testi* (veya kısaca *t-testi*) dir.

William Sealy Gosset *t-testini* ilk 1908 yılında sunmuştur. Gosset İrlanda Dublin’de Guinness Brewery şirketinde üretilen biranın kalitesini kontrol etmekle görevliydi. İşvereni kalite kontrolü için istatistiksel tekniklerin kullanıldığını duyulmasını istemediği için, Gosset geliştirdiği istatistiksel test prosedürünü “Student” mahlasıyla yayımlamıştır. Çünkü Sir Ronald Aylmer Fisher’in öğrencisiydi. Daha sonra, iki grup arasındaki farkı belirtmek için Fisher tarafından kullanılmakta olan *t-değerinin* hesaplanmasını içerdiği için, bu teste Student *t-testi* adı verilmiştir.

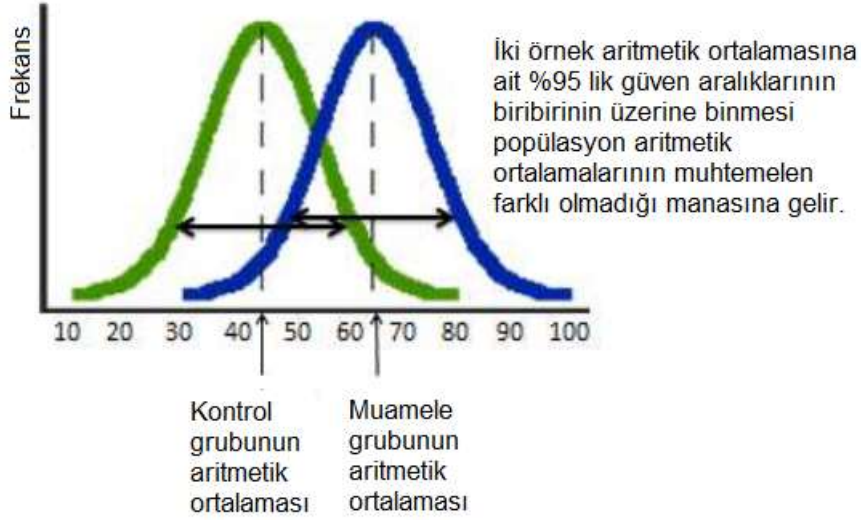
İki grubun aritmetik ortalamalarının birbirlerinden istatistiksel olarak farklı olup olmadığını (yönsüz veya çift-yönlü) veya bir grubun diğer gruba göre istatistiksel olarak daha büyük (veya daha küçük) bir aritmetik ortalamaya sahip olup olmadığını incelemek için *t-testi* kullanılır. Misalimizde, eğer spesiyal matematik müfredatındaki öğrencilerin geleneksel müfredattakilerden daha iyi performans gösterip göstermediklerini incelemeyi arzu ediyorsak, testimiz tek-yönlü olacaktır. Bu hipotez şöyle ifade edilebilir:

$$\begin{aligned} H_0 &= \mu_1 \leq \mu_2 && \text{(sıfır hipotezi)} \\ H_1 &= \mu_1 > \mu_2 && \text{(alternatif hipotez)} \end{aligned}$$

ki burada μ_1 spesiyal müfredata maruz bırakılan öğrencilerin performanslarının popülasyon aritmetik ortalamasını (muamele grubu için) ve μ_2 geleneksel müfredata maruz bırakılan öğrencilerin performanslarının popülasyon aritmetik ortalamasını (kontrol grubu için) temsil etmektedir. Şuna dikkatini çekmek istiyorum ki sıfır hipotezi daima bir “=” işaretine sahiptir ve tüm istatistiksel anlamlılık testlerinin amacı sıfır hipotezini reddetmektir.

Her bir popülasyondan çekilen örneklerin sahip olduğu veriyi kullanarak, popülasyon aritmetik ortalamalarının farkı hakkında nasıl çıkarımda bulunabiliriz? Figür 15.2 deki muamele ve kontrol gruplarına ait skorların hipotetik frekans dağılımlarına bakalım. Kontrol grubu skorları (0-100 lük bir ölçekte) aritmetik ortalaması 45 olan çan-şekilli (normal) bir dağılım gösteriyorken, muamele grubu skorları aritmetik ortalaması 65 olan yine çan-şekilli (normal) bir dağılım gösteriyor. Bu iki aritmetik ortalamanın farklı olduğu görünüyor fakat bunlar gerçekte birer örnek aritmetik ortalaması ($\bar{y}$) ve örnekleme hatasından dolayı, tekabül ettikleri popülasyon aritmetik ortalamalarından (μ) farklı olabilirler. Örnek aritmetik ortalamaları popülasyon aritmetik ortalamalarının belli bir güven aralığındaki olasılıksal kestirimleridir. (Mesela %95 lik güven aralığı örnek aritmetik ortalamasının $\pm$ iki standart hata civarındır ki burada standart hata aynı popülasyondan çekilecek aynı hacimdeki örneklerin aritmetik ortalmaları dağılımının standart sapmasıdır.) Bundan ötürü, örnek aritmetik ortalamaları arasındaki farkın istatistiksel olarak anlamlılığı yalnızca örneklerin aritmetik ortalama skorlarına değil, aynı zamanda örnek aritmetik ortalamalarının frekans dağılımlarının standart hatalarına veya yayılım miktarlarına da bağlıdır. Eğer yayılım genişse (yani, iki çan-şekilli eğri fazlaca çakışyorsa), o zaman iki aritmetik ortalamanın (mesela %95 lik) güven aralıkları da çakışıyor olabilir ve bu taktirde, müteakıl popülasyon aritmetik ortalamalarının farklı olduğu

neticesine varamayız. Mamafih, eğer örnekleme dağılımlarının yayılımları daha dar ise (yani, eğriler azca çakışıyorlarsa), o zaman iki aritmetik ortalamanın güven aralıkları çakışmıyor olabilir ve bu taktirde sıfır hipotezini reddedebiliriz ve deriz ki iki grubun aritmetik ortalamaları arasındaki fark istatistiksel olarak anlamlıdır.



Figür 15.2 Student *t*-testi

Bir *t*-testini yürütmek için, önce iki grubun aritmetik ortalamaları arasındaki farka ait bir *t*-istatistiği hesaplamalıyız. Bu istatistik örnek aritmetik ortalamaları arasındaki farkın gruplardaki skorların değişkenliklerinin farkına (standart hata) oranıdır.

$$t = \frac{\bar{y}_1 - \bar{y}_2}{s_{\bar{y}_1 - \bar{y}_2}}$$

ki burada pay muamele grubu ve (Grup 1) kontrol grubu (Grup 2) arasındaki fark ve payda iki grup arasındaki farkın standart hatasıdır. Standart hata şöyle hesaplanır:

$$s_{\bar{y}_1 - \bar{y}_2} = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

ki burada s^2 ve n her bir grubun sırasıyla varyansı ve hacmidir. Eğer muamele grubunun aritmetik ortalaması kontrol grubunun aritmetik ortalamasından daha büyük ise, *t*-istatistiği pozitiftir. Hesapladığımız *t*-istatistiğinin şans eseri meydana gelmeyecek kadar büyük olup olmadığını incelemek için, *t*-istatistiğinin olasılığına veya *P*-değerine standart ders kitaplarında veya internette bulunan istatistiksel tablolardan bakmalıyız. Veyahutta mesela SAS ve SPSS gibi istatistiksel kompüter yazılımlarını kullanıyorsak, *P*-değeri zaten çıktıda yer alacaktır. Bu değer *t*-istatistiğinin, *t*-testinin tek-yönlü mü yoksa çift-yönlümü olduğunun ve *serbestlik derecesinin* (*df*) bir fonksiyonudur. Serbestlik derecesi bu istatistiğin hesaplanmasında serbestçe değişebilen değerlerin sayısıdır ve her zaman örnek hacminin ve icra edilen testin bir fonksiyonudur. Buna göre, *t*-istatistiğinin serbestlik derecesi şöyle hesaplanır:

$$df = \frac{(s_1^2/n_1 + s_2^2/n_2)^2}{(s_1^2/n_1)^2/(n_1 - 1) + (s_2^2/n_2)^2/(n_2 - 1)}$$

ki bu ifade çoğu zaman $(n_1 + n_2 - 2)$ ye yaklaşır. Eğer bu *P*-değeri arzu edilen bir anlamlılık düzeyinden (diyelim $\alpha = 0.05$) veya bir başka deyişle aslında olmayan bir muamele etkisinin var olduğu neticesine varmak için (Tip I hata) göze aldığımız en yüksek risk düzeyinden (olasılık) daha küçükse, o zaman sıfır hipotezini reddedebiliriz.

Bir muamele grubunun kontrol grubundan anlamlıca daha yüksek bir aritmetik ortalamaya sahip olduğunu gösterdikten sonra, sonraki soru çoğu zaman etki miktarının (ES) veya bir başka deyişle kontrol grubuna göre muamele grubunun etkisinin nispi büyüklüğünün ne olduğudur. Performans skorlarının çıkım değişkeni (y) ve kukla kodlanmış bir muamele değişkeninin tahminci değişken (x) olduğu iki-değişkeli bir GLM ile bir regresyon analizi yaparak, etki miktarını kestirebiliriz. Regresyon doğrusunun eğimi ($\beta_1 = \Delta y / \Delta x$) ve aynı zamanda muamele değişkeninin regresyon katsayısı (β_1) etki miktarının bir kestirimidir. Yukarıdaki misalde, x iki değer (0 ve 1) alabilen bir kukla değişken olduğu için, $\Delta x = 1 - 0 = 1$ olduğundan ötürü, etki miktarı veya β_1 basitçe muamele ve kontrol gruplarının aritmetik ortalamaları arasındaki farktır ($\Delta y = y_1 - y_2$).

Faktöriyel Dizaynlar

Bir önceki misali genişleterek, diyelim ki spezial müfredatın (muamele) etkisi ve geleneksel müfredatın (kontrol) öğrenci performans skorlarına etkisi öğretim süresine (haftada 3 veya 6 saat) bağlıdır. Şimdi, müfredat tipi (spezial veya geleneksel) ve öğretim süresi (haftada 3 veya 6 saat) olmak üzere iki faktörümüz var ve böylece 2×2 faktöriyel bir dizayna sahibiz. Böyle bir dizayn yalnızca *ana etkiler* adı verilen her bir faktörün bağımsız etkisini kestirmemize yardımcı olmakla kalmaz, aynı zamanda *interaksiyon etkisi* adı verilen her iki faktörün ortak etkisini de kestirmemize olanak sağlar. Bu iki-yönlü faktöriyel dizayn için genel doğrusal model şöyle yazılır:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_1 x_2 + \varepsilon$$

ki burada y öğrencilerin deney-sonu performans skorlarını temsil ederken, x_1 müfredat tipi (spezial veya geleneksel) ve x_2 öğretim süresidir (haftada 3 veya 6 saat). Dikkat etmelisin ki hem x_1 hem de x_2 kukla değişkenlerdir ve her ne kadar bir oran-ölçeği değişkeni gibi (3 veya 6) görünüyor olsa da, x_2 bu faktöriyel dizaynda iki kategoriye temsil etmektedir. Regresyon katsayıları β_1 ve β_2 ana etkilere ait etki miktarı kestirimlerini verirken, β_3 interaksiyon etkisine ait etki miktarı kestirimini verir. Alternatif olarak, aynı modeli iki-yönlü bir ANAOVA ile analiz edebiliriz. Çoklu tahminci değişkenleri havi regresyon analizi bezen *çoklu regresyon* olarak adlandırılır, öyle ki çoklu çıkım değişkenlerini havi *çok-değişkeli regresyon analizinden* farklıdır.

İnteraksiyon etkilerinin interpretasyonu hakkında şunu not edelim. Eğer β_3 anlamlıysa, bunun manası şudur: Öğrencilerin performans skoruna, müfredat tipinin etkisi öğretim süresine bağlıdır. Bu halde, müfredat tipinin veya öğretim süresinin bağımsız etkilerinin (sırasıyla β_1 ve β_2) manalı interpretasyonlarını yapabilmeyiz. Çünkü bu iki etki birbirinden ayrılabilir. Ana etkilerin interpretasyonları yalnızca interaksiyon etkisi anlamsız çıktığı zaman tek başlarına yapılabilir.

Kovaryeteler bir faktöriyel dizayna yeni regresyon katsayıları ile (mesela β_4) yeni değişkenler olarak dahil edilebilir. İlgilenilen tahminci değişkenler kukla değişkenlerle ifade edilmiş olsalar dahi, kovaryeteler aralık veya oran ölçeğinde ölçülebilirler. Kovaryetelerin interpretasyonları diğer tahminci değişkenleriyle aynı kurallara tabidir.

Diğer Nicel Analizler

Genel veya genelleştirilmiş doğrusal modellere dayanan kullanışlı pek çok istatistiksel teknik vardır. Bunlardan bazılarını burada kısaca değineceğiz. İlgilenen okuyucular bu teknikler hakkında daha fazla enformasyon için ileri düzeydeki ders kitaplarına veya istatistik derslerine

başvurabilirler.

- *Faktör analizi* bir veri indirgeme tekniğidir öyle ki gözlenen çok sayıdaki ölçüyü (maddeyi), bunların altta-yatan iki-değişkeli korelasyon biçimlerine dayanarak, faktör adı verilen daha az sayıdaki gözlenmeyen (latent) değişkenler içerisinde istatistiksel olarak toplandırmak için kullanılır. Sosyal bilim araştırmasında bu teknik çok-maddeli ölçeklerde yakınsak geçerliliğin ve ayrımsal geçerliliğin değerlendirilmesinde yaygınca kullanılır.
- *Diskriminant analzi* bir sınıflandırma tekniğidir öyle ki verilen bir gözlemi, tahminci değişkenlerin bir doğrusal kombinasyonuna dayanarak, bir kaç nominal kategoriden birine yerleştirmeyi hedefler. Bu teknik çoklu regresyona benzer fakat şu farkla ki yanıt değişkeni niteldir (yani adlandırma ölçeğinde). Mesela geniş çaplı yoklamalarda teşhis edilen belirgin vasıflarına dayanarak müşterileri veya ürünleri kategorilerde sınıflandırmak gibi pazar uygulamalarında popülerdir.
- *Lojistik regresyon* (veya lojit modeli) bir GLM dir öyle ki çıkım değişkeni iki-değerlidir (0 veya 1) ve başarılı çıkımın olasılığı bir lojit transformasyonu vasıtasıyla yanıt değişkenlerinin bir kümesine ait veriden kestirilir. Bir misali yaş, vücut kütle endeksi, egzesiz rejimi ve benzeri tahminci değişkenlere dayanmak suretiyle, spesifik bir periyotta kalp krizi geçirme olasılığının tahmin edilmesidir. Lojistik regresyon medikal bilimlerde aşırı popülerdir. Etki miktarı kestirimi bir ihtimal oranına (*odds ratio*) dayanır ve bir olayın bir grupta meydana gelmesinin, diğer gruba karşı, ihtimalini temsil eder.
- *Probit regresyon* (veya probit modeli) bir GLM dir öyle ki çıkım değişkeni 0 ve 1 arasında değişebilir (veya 0 ve 1 kesikli değerlerini alabilir) ve standart normal dağılıma uyduğu kabul edilir. Regresyonun amacı çıkım olasılığını tahmin etmektir. Sigortacılıkta, finansal hizmetlerde ve başka sektörlerde, prediktif analiz için popüler bir tekniktir. Mesela kredi başvurusunda bulunan bir kişinin kredi reytingine, kazancına, borcuna ve daha başka enformasyonuna dayanarak bir kredi skoru hesaplanır. Probit ve lojit regresyon aynı veri uygulamalarında benzer regresyon katsayıları üretme eğilimi gösterirler. Ancak, lojit modelinin hesaplanması ve interpretasyonu daha kolaydır.
- *Path analizi* bir değişkenler kümesindeki yönlü ilişkileri analiz etmek için kullanılan bir çok-değişkeli GLM tekniğidir. Bir eşitlikteki yanıt değişkeninin bir diğer eşitlikte açıklayıcı değişken olduğu kompleks nomolojik modellerin incelenmesine olanak sağlar ve çağdaş sosyal bilim araştırmasında yaygınca kullanılır.
- *Zaman serisi analizi* bir zaman periyodu boyunca değişen zaman serisi verisini analiz etmekte kullanılan bir tekniktir. Uygulamadaki misalleri arasında borsa dalgalanmalarını ve kentsel suç oranlarını tahmin etmeyi sayabiliriz. Bu teknik ekonometride, matematiksel finansta ve sinyal işlemede popülerdir. Aynı değişkenin zaman boyunca değişen değerleri içerisindeki korelasyonu ifade eden oto-korelasyon için spesiyal teknikler kullanılır.

Bölüm 16

Araştırma Etiği

Sözlükte (Webster's Dictionary) etik belli bir meslek veya grup dahilindeki davranışlara ait standartlara uygunluk olarak tanımlanır. Böyle standartlar mesleki davranış kodları olmakla beraber, disiplinler düzeyinde tanımlanırlar ve bazen üniversitelerde ilgili kurullar tarafından gözetilirler. Açıkça belirtilmiş olmasa dahi, bilimcilerin ait oldukları bilimsel topluluk tarafından paylaşılan genel anlaşmaların farkında olmaları ve onlara uymaları beklenir. Bu genel anlaşmalar bilimin mesleki olarak icrasında kabul edilebilir ve kabul edilebilmez davranışları vazeder. Mesela, bilimciler veri toplama, analiz ve interpretasyon prosedürlerinde bilimin veya bilimsel metodun prensipleriyle çelişecek şekilde manipülasyonlar yapmamalıdır veya kişisel çıkarlarını öne almamalıdır.

Araştırma etiği niçin önemlidir? Çünkü, bilim kendi kişisel çıkarlarını öne alan ve bilimsel tahkikatın normlarına ters düşen faaliyetlere angaje olan kişiler veya organizasyonlar tarafından etik-olmayan yollarla sıkça manipüle edilmektedir. Ecza devi Merck'in Vioxx davası bunun klasik bir misalidir. Ki bu şirket Vioxx adlı ilacın ölümcül yan etkilerini bilim camiasından gizlemişti ve bu da bu ilacı kullananların 3468'inin çoğunlukla kalp krizinden dolayı ölümüyle sonuçlanmıştı. Yıl 2010 da, bu şirket 4.85 milyar Amerikan Doları ceza ödemenin yanı sıra şirketin kendi ilaç geliştirme prosesinin emniyetini göz altında tutmak üzere iki bağımız komite ve yetkili bir medikal memur atanmasını kabul etmişti. Merck'in davranışı etik değildi ve veri toplamaya, analize ve interpretasyona dair bilimsel prensipleri ihlal etmişti.

Etik doğru ve yanlış arasındaki ahlaki ayrımdır ve etik olmayan bir şey illa ki yasadışı olmayabilir. Eğer bir bilimcinin davranışı etik ve yasa arasında bir gri bölgeye düşerse, yasa nezdinde suçlanmayabilir fakat yine de kendi meslek topluluğunca dışlanabilir, mesleki itibarı zedelenabilir ve hatta mesleğinin kusurlu icrası bakımından işinden olabilir. Bu etik normlar bir topluluktan diğerine farklılık gösterebilir. Bizim burada referans olarak alacağımız etik standartlar Batı ülkelerinde bilimsel araştırma için uygulananlar olacak.

Bilimsel Araştırmanın Etik Prensipleri

Bilim camiasında geniş ölçüde kabul edilen etik davranışa ait ilkeler aşağıdaki üzeredir.

Gönüllü katılım ve zararsızlık. Bir araştırma projesinin denekleri öğrenme-çalışmasına gönüllü katılımlarının, öğrenme-çalışmasından herhangi bir zamanda maddi veya manevi herhangi bir bedel ödmeden çıkabilme özgürlüklerinin ve projeye katılmalarının veya katılmamalarının sonucunda bir zarar görmeyeceklerinin farkında olmalıdırlar. Gönüllü katılım prensibinin en rezil ihlalleri muhtemelen İkinci Dünya Savaşı sırasında Nazi araştırmacılarının hapishanelerde yürüttükleri zoraki medikal deneylerdir. Ki bunlar savaş sonrası Nuremberg Duruşmalarında belgelenmişlerdir. (Bu deneyler aynı zamanda “insanlığa karşı işlenen suç” teriminin orijini.) Daha az bilinen ihlallerden biri Birleşik Devletler Halk Sağlığı Servisi tarafından 1932-1972 yılları arasında yürütülen Tuskegee frengi deneyleridir. Ki bu deneylerde frengiye yakalanmış yaklaşık 400 yoksul Afrikalı-Amerikalı adama mesela spinal doku gibi yanlış tedaviler uygulanmıştır. Oysa, frenginin efektif bir tedavisi olarak, penisilin o tarihte kabul edilmişti. Denekler ölümcül bir tehdide maruz kalmamış olsalar bile, katılımlarının sonucunda ızdırap çekmemelilerdi. Yıl 1971 de, psikolog Philip Zimbardo *Stanford Prison Deneyini* başlattı ki bu deneyde Stanfoddan denek olarak devşirilen öğrencilere rastgele mahpus ve gardiyan gibi roller verildi. Mahpus öğrencilerin rol icabı mahpuslukları sonucunda

psikolojik zarara uğradıkları, gardiyan rolündeki öğrencilerin sadizm sergiledikleri ve bunların daha sonra kendi öz-imajlarını tehdit edebileceği ortaya çıkınca, deney sonlandırıldı.

Bugün, eğer bir öğretmen öğrencilerinden bir anketi doldurmalarını isterse ve katılımlarının gönüllü olduğunu bildirirse, öğrenciler ankete katılmamalarının o dersten alacakları notu herhangi bir şekilde düşürebileceğinden korkmamalıdır. Mesela, katılanlara bonus puanlar vermek ve katılmayanlara vermemek etik değildir çünkü bu katılmayanlar için bir dezavantaj yaratır. Böyle şartlardan kaçınmak için, öğretmen katılmayanlar için alternatif bir görev verebilir öyle ki onlar da ankete katılmadan bonus puanları alabilsin. Ya da ankete katılsın veya katılmasın her öğrenciye bonus puan verebilir. Dahası, öğrenme-çalışmasının yanıtları kaydedilmeden önce, tüm katılımcılara birer **muafakat (/informed consent)** formu verilmeli ve imzalatılmalıdır. Bu formda, onlara katılmama hakları ve çekilme hakları açıkça betimlemelidir. Bir medikal öğrenme-çalışmasında, bu formda aynı zamanda katılımlarından dolayı deneklerin maruz kalabilecekleri olanaklı riskler de belirtilmelidir. Yaşı 18 in altında olan denekler için, bu form ebeveyni veya yasal vasileri tarafından imzalanmalıdır. Araştırmacılar veri toplama prosesinin tamamlanmasından sonra kendi disiplinlerine veya işyerlerine ait bilim yapma normlarına uymak için bu muafakat formlarını bir zaman periyodu boyunca (çoğunlukla 3 yıl) saklamalıdır.

Anonimlik ve mahremiyet. Deneklerin çıkarlarını ve müstakbel esenliklerini korumak için, bir öğrenme-çalışmasında kimlikleri saklı tutulmalıdır. Bunu sağlamak için anonimlik ve mahremiyet prensipleri kullanılır. **Anonimlik** araştırmacının veya araştırma raporunu veya makaleyi okuyanların spesifik bir yanıtlayıcının verdiği yanıtı teşhis edebilmemesi demektir. Bilimsel araştırmada anonimliğin bir misali yoklamayı kimin yanıtlayıp kimin yanıtlamadığını izlemek için kimlik numaralarının kullanılmadığı bir posta yoklamasıdır. Mesela öğrencilerin madde kullanımı veya yasadışı müzik indirmeleri gibi sapkın veya arzu edilmeyen davranışlara dair öğrenme-çalışmalarında, eğer deneklerin anonimlikleri garanti edilmezse, hakiki yanıtlar elde edilemez. Dahası, anonimlik deneklerin yasal yaptırım veya gelecekte kendilerini teşhis edebilecek veya takip edebilecek başka otoriteler karşısında masuniyetlerini garanti eder.

Mesela yüz-yüze görüşmeler gibi bazı araştırma dizaynlarında, anonimlik olanaklı değildir. Mesela boylamsal saha yoklamaları gibi başka dizaynlarda, anonimlik arzu edilmez çünkü araştırmacının boylamsal veri analizi için aynı denekler üzerinde zamanın farklı noktalarında alacağı yanıtları eşleştirmesine mani olur. Böylesi şartlar altında, deneklerin **mahremiyetleri** garanti edilmelidir. Yani araştırmacı deneklerin yanıtlarını teşhis edebilmeli fakat onların kimliklerini herhangi bir raporda, makalede veya halka açık forumda açığa vurmaya söz vermelidir. Mahremiyetin koruyuculuğu anonimliğe göre daha zayıftır çünkü sosyal araştırma verisi mesela Birleşik Devletler mahkemelerinde olduğu gibi çoğu ülke hukukunda “kişisel iletişim” kapsamında değildir. Mesela, Exxon Valdez süpertankeri Alaska’da Valdez limanı yakınında 10 milyon ham petrol varilini denize döktükten iki yıl sonra, ekonomik ve çevresel zarara uğrayan topluluklar ailelerin artan psikolojik problemlerine dair kişisel ve utandırıcı detaylar hakkında hanehalkı yoklamaları yapmak için San Diegoda bir araştırma firmasıyla anlaştılar. Çoğu Yerli Amerikalılar için, kültürel normlarından dolayı, böylesi halka açık ifadelerde bulunmak bilhassa acı verici ve zor olduğu için, yanıtlayıcılara yanıtlarının mahremiyeti garanti edildi. Toplanan kanıt mahkemeye sunulduğu vakit, verdikleri cevaplara nazaran yanıtlayıcıları çapraz-inceleme için, Exxon mahkemeden orjinal yoklama anketlerinin (kimlik bilgileri ile) celbini talep etti. Görüşmecilerin onlara mahremiyet sözü vermiş olmalarına rağmen, mahkeme bu talebi uygun buldu. Allahtan, kurbanlar açık duruşmada ifade vermeye zorlanmadan önce, Exxon Valdez davası anlaşmayla sonuçlandı. Fakat mahremiyetin benzer şekilde ihlal edilmesi potansiyeli hala mevcut.

Aşırı bir vaka olarak, Washington State Üniversitesinde bir lisansüstü öğrencisi olan Rick

Scarce hayvan hakları aktivistleri ile katılımcı gözlem öğrenme-çalışmaları yürüttü ve bulgularını 19901 yılında *Ecowarriors: Understanding the Radical Environmental Movement* adlı bir kitapta yazdı. Yıl 1993 te, öğrenmeye-çalıştığı aktivistleri teşhis etmek üzere, Scarce bir büyük jüri tarafından çağrıldı. Araştırmacı, Amerikan Psikoloji Derneğinin bir üyesi olmak suretiyle, etik yükümlülüklerine sadık kalarak, büyük jürinin sorularına cevap vermeyi reddetti ve Spokane Bölge Hapishanesinde 159 gün geçirmek zorunda kaldı. Rick Scarce'ninkine benzer sıkıntılar çekmemek için, lüzümü kalmayınca, araştırmacılar belgelerden ve veri dosyalarından tüm kimlik enformasyonunu silmeliler. Yıl 2002 de, Birleşik Devletler Sağlık ve İnsani Hizmetler Bölümü, araştırma projelerine katılanları polisten ve diğer otoritelerden korumak için, bir "mahremiyet sertifikası" tebliğ etti. Her araştırma projesi bu koruma kapsamına girmez, fakat çoğu durumda katılımcıların mahremiyetini korumak için önemli bir destek temin edebilir.

Açığa-vurma. Çoğunlukla, araştırmacılar, öğrenme-çalışmasına katılıp katılmayacaklarına karar vermelerine yardımcı olmak üzere, veri toplamadan önce, araştırmaları hakkında potansiyel deneklere bir miktar enformasyon sağlamakla yükümlülerdir. Mesela, öğrenme-çalışmasını kimin yürüttüğü, maksadının ne olduğu, hangi çıkımların beklendiği ve sonuçların kime yararı olacağı gibi. Mamafih, bazı durumlarda, böylesi enformasyonu açığa vurmamak deneklerin yanıtlarını potansiyel olarak yanıltılabılır. Mesela, eğer bir öğrenme-çalışmasının maksadı deneklerin "grup-düşüncesine" riayet etmek için ne ölçüde kendi görüşlerini terke edeceklerini incelemekse ve denekler önce bir mevzuda başkalarının görüşlerini dinledikten sonra kendi görüşlerini söyleyecekleri bir deneye katılacaklarsa, o zaman deneyden önce öğrenme-çalışmasının maksadını açığa vurmamak denekleri muameleye karşı duyarlılaştıracaktır. Böylesi şartlar altında, öğrenme-çalışmasının maksadı öğrenme-çalışmasından önce açığa vurulabilmesede dahi, veri toplama prosesinin bitmesinden hemen sonra kısaca katılımcılara deney sırasında maruz kaldıkları potansiyel riskler veya zarar açıklanabilir.

Analiz ve raporlama. Araştırmacıların aynı zamanda bilimsel camiaya karşı öğrenme-çalışmalarında verinin nasıl analiz edildiğine ve sonuçların nasıl rapor edildiğine dair etik yükümlülükleri vardır. Beklenmeyen veya negatif bulgular, her ne kadar araştırma dizaynına veya bulgulara gölge düşürüyor olsa bile, tamamıyla açığa vurulmalıdır. Benzerce, bir öğrenme-çalışması tamamlandıktan sonra, şans eseri veya veri madenciliği vasıtasıyla pek çok ilginç ilişkiler keşfedilebilir. Böyle bulguları tasarlanan dizaynın bir ürünüymüş gibi sunmak ta yine etik değildir. Başka bir deyişle, pozitivist araştırmada hipotezler veri analizinin sonuçlarına dayanan olgulardan sonra inşa edilmemelidir, çünkü pozitivist araştırmada verinin rolü zaten hipotezleri est etmektir, onları inşa etmek değil. İlgilendiği hipotezi ispatlamak veya çürütmek için veya farklı veri kümelerinden daha çok sayıda makale üretmek için, araştırmacının verisini farklı segmentlere "dilimlemesi" de yine etik değildir. Kısmi, eksik veya münasip-olmayan veri analizine dayanan sorgulanabilir iddiaları geçerliymiş gibi sunmak ta yine namussuzluktur. Bilim açıklıkla ve dürüstlikle ilerler ve araştırmacılar, araştırmalarındaki problemleri tamamen açığa vurarak, bilime ve bilimsel camiaya en iyi şekilde hizmet edebilirler, öyle ki başka araştırmacılar benzer problemlerden kendilerini koruyabilsinler.

Kurumsal Gözetim Kurulları

İnsan denekleri ihtiva eden öğrenme-çalışmalarında araştırma etiği, Birleşik Devletler'de, federal yasa tarafından idare edilir. Kendi araştırma projelerini desteklemek için federal fonlara başvurmak isteyen mesela üniversite veya hastahane gibi kurumlar veya kuruluşlar insan deneklerin haklarını ve korunmalarını idare eden federal yasalara uyduklarını tesis etmelidirler. Bu proses her kurumda veya kuruluştaki Kurumsal Gözetim Kurulu (IRB) adı verilen bir uzmanlar kurulu tarafından yürütülür. Gönüllü katılım, zararsızlık, anonimlik, mahremiyet ve

benzeri prensiplerin yerine getirildiğinden ve insan deneklerin maruz kalacağı risklerin minimize edildiğinden emin olmak için, IRB insan denekleri ihtiva eden tüm araştırma önerilerini gözden geçirir. Federal yasalar spesifik olarak her ne kadar federal fonların desteklediği projeler için uygulanıyor olsa bile, aynı standartlar ve prosedürler fon desteği almayan projelere veya hatta öğrenci projelerine dahi uygulanır.

IRB tasvip prosesi strüktürlü bir başvurunun tamamlanmasını gerektirir. Bu proseste araştırma projesi hakkında eksiksiz enformasyon, araştırmacılar (başlıca soruşturmacılar) ve deneklerin haklarının nasıl korunacağına ait detaylar temin edilir. Mesela muvafakat formu, araştırma anketi veya görüşme protokolü gibi ilave belgelere de ihtiyaç duyulabilir. Araştırmacılar yine etik araştırma prensiplerine aşina olduklarını da göstermelidirler. Bunun için bir araştırma etiği kursuna katılım sertifikası temin edebilirler. Veri toplama ancak söz konusu proje IRB tarafından tasvip edildikten sonra başlatılabilir.

Mesleki Etik Kodu

Araştırmacılara ait çoğu meslek derneği kendi üyelerinin kabul-edilebilir ve kabul-edilebilirmez mesleki davranışlarının hangileri olduğunu betimleyen resmi etik kodları tesis etmişlerdir. Bir misal olarak, enformasyon sistemleri alanındaki araştırmacıların küresel meslek örgütü olan Enformasyon Sistemleri Derneği'nin (AIS) etik kodu Tablo 16.1 de özetlenmiştir. (Bu etik kodun tamamı anılan derneğin web sitesinde bulunabilir.) Benzer etik kodları başka disiplinler için de yine mevcuttur.

AIS etik kodu etik ihlallerini iki kategoride gruplandırıyor. Kategori I mesela intihal (/plagiarism) ve veriyle, araştırma prosedürleriyle veya veri analiziyle oynamak gibi kabahatleri kapsıyor ve dernekten çıkartılmaya, işten atılmaya veya yasal yaptırımlara kadar cezası olabiliyor. Buna tevessül eden araştırmacılar mesleki itibarları bakımından yıkıcı zarara uğrayabilirler. Kategori II mesela araştırma deneklerinin haklarına riayet etmemek, araştırma projelerinin orjinallliğini saptırmak ve başkalarının yayımladığı veriyi haber vermeden kullanmak gibi daha az ciddi kabahatleri kapsıyor. Bunları işlemek te yine araştırmacının mesleki itibarına zarar verebilir, dergilerinin yayımlarına maruz kalabilir ve benzeri. Bu kod aynı zamanda iyi araştırma davranışları, etik kabahatler belirlendiği zaman ne yapılacağı (hem kabahatli hem kurban için) ve etik ihlali vakalarının ele alınmasında AIS tarafından icra edilecek proses hakkında kılavuzluk ediyor. Bunun gibi kodlar elbette etik olmayan davranış tamamen bertaraf etmiyor fakat bilimsel camiada etik davranışın sınırlarının belirlenmesine ve etik kabahatlerin işlenmesinin azaltılmasına kesinlikle yardımcı oluyor.

KATEGORİ BİR: Bu kategorideki kodlara DAİMA riayet edilmelidir. Bu kodların dikkate alınmaması halinde, ciddi etik ihlalleri ortaya çıkar. Ciddi ihlaller sizin akademik derneklerden tard edilmenize, işinizden atılmanıza, size karşı yasal tedbirlerin alınmasına ve akademik itibarınızın potansiyel olarak ölümcül yaralar almasına neden olabilir. 1. İntihal yapmayın. 2. Veriyi, araştırma prosedürlerini veya veri analizini uydurmayın veya tahrif etmeyin.
KATEGORİ İKİ: Bu kategorideki kod tavsiye edilen etik davranıştır. Bunları veya daha başka mesleki adabı dikkate almamak, daha az ciddi olmakla beraber, itibarınıza zarar verebilir, editoryal yaptırımlara maruz bırakılabilirsiniz, camiada ayıplanabilirsiniz, yasal süreçlere tabi tutulabilirsiniz ve meslektaşlarınızın size karşı iyi niyetini kaybedebilirsiniz. 3. Araştırma deneklerinin haklarına saygı gösterin. Bilhassa onların enformasyon mahremiyeti haklarına ve angaje olmaları istenen araştırmanın ve aktivitelerin mahiyeti hakkında bilgilendirilme haklarına. 4. Sunduğunuz makalelerin orijinalliği hakkında editörleri ve konferans programı başkanlarını yanıltmayın. 5. Size tevdi edilen editörlük, hakemlik veya danışmanlık görevlerinizin getirdiği otoriteyi ve sorumluluğu istismar etmeyin ve kişisel ilişkilerinizin yargınıza müdahil olmadığından emin olun. 6. Makale, proje veya yazılım hakemliği yaparken veya dış kaynaklardan iş alırken, objektif ve tarafsız davranabilme kabiliyetinizi zaafa uğratabilecek herhangi bir somut çıkar çatışmasını declare edin. 7. Başkalarına ait yayımlanmış veriyi ikrar etmeksizin veya yayımlanmamış veriyi hem izin almaksızın hem de ikrar etmeksizin alma veya kullanmayın. 8. İster meslektaşın olsun ister öğrencilerin olsun, araştırmaya katılanların tümünün sübstansif katkılarını, entellektüel katkılarını göre ikrar edin. 9. Mesela hakemlik gibi bir proses mucibince görebileceğin başkalarının yayımlanmamış yazılarını, eformasyonunu, fikirlerini, kavramlarını veya verisini yazarının iznini almaksızın kullanmayın. 10. Arşiv materyalini ancak arşiv kaynağının kurallarına uyarak kullanın. ÖĞÜT: Kendinizi yazarlık çekişmelerinden, yanlışlardan, kabahatlerden ve hatta yasal yaptırımdan koruyabilmeniz için bazı fikirler. 11. Bağlantılı olduğunuz her akademik çalışma için original yazarlığınızı geçerlemek için lazım olan belgeleri ve veriyi muhafaza edin. 12. Yeni bir entellektüel katkı sağlayacak olsa dahi, kendinize ait eski fikirleri yeniden yayımlamayın. 13. Veri kümesi sahipliği meselelerini veri toplama işi bitmeden önce halledin. 14. Eğer şüpheye düşerseniz, mütenasip meslektaşlarınıza danışın.

Tablo 16.1. Enformasyon Sistemleri Derneği'nin Etik Kodu

Bir Etik İhtilafı

Robert Allen “Laud” Humpreys Amerikalı bir sosyolog ve yazardır ki 1970 yılında yayımlanan *Tearoom Trade* adlı tezi ile çok tanınmıştır. Bu kitap parklardaki umumi tuvaletlerde anonim erkek homoseksüel buluşmaların etnografik bir muhasebesidir. Bu teamül Birleşik Devletler homoseksüel argosunda “tea-rooming” olarak bilinir. *Tearoom* aktivitelerine katılanların çoğunluğunun toplum içerisinde geleneksel aile hayatı yaşayan görünüşte heteroseksüel adamlar olmaları olgusu Humphreys’in merakını kamçılıdır. Mamafih, *tearoom* ziyaretleri sırasında anonimliklerini muhafaza etmek bu adamlar için önemliydi.

Tipik olarak, *tearoom* buluşmalarında üç kişi oluyordu. İki erkek cinsel ilişkiye girerken “watchqueen” adı verilen bir kişi gözcülük yapıyordu. *Watchqueen* ilişkiye giren iki erkeğin eylemini bir voyör olarak izleyerek zevk alırken, aynı zamanda polise veya başka kişilere karşı onları uyarmakla vazifeliydi. Bu deneklere başka türlü ulaşmak mümkün olmadığından, Humphreys bir *watchqueen* taklidi yaparak umumi tuvaletlerde kendini gösterdi. Bir katılımcı gözlemci olarak, tıpkı politik protestoları veya başka herhangi sosyal fenomenleri öğrenmeye-çalışırken yapılabildiği gibi, Humphreys tezi için saha gözlemleri yapabilmeyi başardı.

Humphreys deneklerinden daha çok enformasyon elde etme ihtiyacı duydu. Fakat katılımcıların görüşmeye ve kimliklerini açığa vurmaya isteksiz olmalarından dolayı, Humphreys katılımcıların araçlarının plakalarını imkan bulduğunca kaydetti ve adlarına ve adreslerine umumi veritabanlarından ulaştı. Daha sonra, tanınmamak için kılık değiştirerek bu adamları evlerinde ziyaret etti ve bir yoklama yürüttüğünü söyleyerek, başka yoldan elde edemeyeceği kişisel veriyi topladı.

Humphreys’in araştırması bilimsel camiada dikkate şayan ihtilaf yarattı. Eleştirenlerin bir kısmı dediler ki bilim adına başkalarının hususiyetine (yani özel hayatına) tecavüz etmemeliydi. Bir kısmı da açıkça art niyetli olarak, katılımcıları sadece bir *watchqueen* olduğuna inandırmaktaki hileli davranışına dair endişelerini belirttiler. Katılımcıların umumi tesisleri kullanıyor olmaları hasebiyle, *tearoom* aktivitelerini gözlemeyi kabul-edilebilir bulanlar dahi katılımcıların evlerinde kılık değiştirerek takip yoklaması yapmanın etik olmadığını düşündüler. Çünkü onların ev adreslerini ele geçirme yolu ve muvafakatlarını almaması etik değildi. Çok az araştırmacı Humphreys’in yaklaşımını haklı buldu ve dediler ki bu veri toplamak için başka bir yolu olmayan, soruşturmaya değer önemli bir sosyolojik fenomendir ve Humphreys deneklerinin kimliklerini hiç kimseye açık etmediği için, yapılan hilenin bir zararı yoktur. Bu ihtilaf şimdiye kadar çözülmedi ve araştırma etiği sınıflarında ve forumlarında hala ateşlice tartışılmaya devam etmektedir.

Appendix

Sample Syllabus for a Doctoral Course

QMB 7565: Introduction to Research Methods University of South Florida College of Business Fall Semester 2011
--

Objectives:

The purpose of this course is to introduce doctoral students to the process of conducting academic research. We will learn about how to think and act like a researcher in conceptualizing, designing, executing, and evaluating “scientific” research projects. Part of the class will require you to design and write an independent research proposal (with your professor’s help). In addition, the finals exam will prepare you for the research methods section of your Ph.D. comprehensives exam.

Structure:

This class is designed in a seminar format. The heart of any Ph.D. seminar is discussion and analysis of assigned readings. To do that, you must read all assigned readings before class, think about these issues throughout the semester, debate these issues with your classmates, and synthesize these issues mentally to develop yourself as a researcher. Please note that if you do not come to class fully prepared, you will be totally lost and will not learn anything from the class.

Being a doctoral seminar, this course will entail: (1) a much higher workload than any Masters-level course you have encountered thus far, (2) a heavy dose of boring readings, and (3) a substantial amount of critical and often frustrating thinking. This is not an easy class and you will not receive an easy grade. If you are challenged by the demanding nature of this class, you should drop this class and drop out of the Ph.D. program.

Books and Materials:

Bhattacharjee, A. *Social Science Research: Principles, Methods, and Practice*, Ver 2.0, 2011, Free download from Blackboard (my.usf.edu).

Kuhn, T. J. *The Structure of Scientific Revolutions*, University of Chicago Press, Chicago, 1996. This entire book must be read prior to Week 4. This \$7.50 paperback (Amazon price) looks deceptively small, but is actually a heavy drill, so start reading it right away.

Papers: Download from Blackboard (my.usf.edu). See Class Schedule.

Grading:

Grade components:	Finals exam	35 points
	Research proposal	35 points
	Paper review	10 points
	Class participation	20 points

Grading scale: A+: 97-100; A: 92-97; A-: 90-92; B+: 88-90; B: 82-88; B-: 80-82; C+: 78-80; C: 70-78

Finals Exam:

This is a 3-hour, comprehensive exam, in which you will be asked 3-4 multi-part essay-type questions, similar to the ones in your Ph.D. comprehensive exams. Sample exam questions can be downloaded from Blackboard. You can either hand-write your answers or type them on a computer. Your answers should demonstrate (a) synthesis of materials covered in class and (b) your own critical analysis of these materials, and not a mere regurgitation of the papers or the professor's comments. You will be graded solely on quality of your answer and not on how many pages you write. You will lose credit if your answer is irrelevant (of the rambling type), does not contain enough details (too general or vague), or demonstrate just a superficial level of understanding of the materials discussed in class. Exam is open-book/open-notes, but please be forewarned that having all the materials in front of you will not help, if you didn't bother to prepare for the exams ahead of time. Given the avalanche of topics and materials covered in class, cramming the week before the exams will be too little and too late.

Research Proposal:

You will apply your learning about the research process into a research proposal that is due at the end of the semester. This proposal will be 10-12 pages single-spaced, excluding references and appendices. You can select any research problem of your choice for this proposal, but preferably in the positivist tradition of inquiry. However, the proposed research must be original (i.e., not something you are doing for another class or another professor), must examine a real problem (i.e., not a hypothetical or "toy" problem), and must be of at least a conference-level quality. Your proposal must include five sections: (1) research problem and significance, (2) literature review, (3) theory and hypotheses, (4) research methods, and (5) research plan. Data collection or analysis is not necessary. But if you want to do an interpretive research project instead, then some data collection and analysis will be needed (and the above structure will also change – talk to me about these changes). Project deliverables are due throughout the course of the semester, as we cover corresponding topics in class. This will allow me to give you early feedback and correct problems well before the final due date. Note that intermediate deliverables are not graded, only your final proposal will be graded. You will also present your final proposal during the last week of class in a 15-20 minute formal

presentation (plus a 5 minute question/answer session), as you would typically do at an academic conference.

Paper Review:

A key component of an academic career is critically evaluating others' research. During the second half of the semester, you'll be asked to write a formal review (critique) of a paper submitted for publication to a leading business journal. To help you prepare, I'll discuss in class how to write reviews, give you a framework for review, and have you do one practice round of review in class for an actual journal submission. After you write your review, I will give you actual reviews by anonymous reviewers/editors for this paper so that you can compare your own review with that of professional reviewers, and see what you missed. The graded paper review is take-home, and due one week before the finals. You may need to do some background research or read some additional papers before writing your review report, but you must not collaborate or discuss the paper with anyone in or out of class.

Class Participation:

Each paper discussed in class will be assigned a "lead discussant," who will be responsible for (1) preparing a one-page synopsis of those papers (in a structured format) and (2) leading class discussion on that paper. You will receive full class participation grade if you turn in synopsis of all assigned papers on time and do a reasonable job in class discussion. Please see the sample synopsis to get an idea of how to structure these synopses. These synopses can be useful study aids for the exams, but only if you do a thorough job with them. Bring enough copies of the synopses to share with the rest of the class. I'll give you immediate feedback on your synopsis as we discuss the paper in class, so that you can try to improve your synopsis next time. If you have to miss a class or would like to do a different paper, you can exchange your assigned paper with a classmate. However, if there is a miscommunication and no synopsis is available for a paper on the discussion date, the student originally assigned that paper will be penalized in his/her class participation grade. Irrespective of synopsis assignment, everyone is expected to attend all classes, read all papers, and contribute to all class discussions.

Class Policies:

Attendance: I do not take formal attendance, but I do keep track of who is coming to class and who is not. If you think you will miss more than one week of class during the semester, you should drop this class.

Academic honesty: Plagiarism in any form is banned and will result in a straight FF grade. Please refer to USF's academic honesty policy in your student handbook.

Disability: Students requiring disability accommodations should notify me within the first two weeks of class, with a letter from the Student Disability Services Office.

Cell phones: Cell phones must be turned OFF during the duration of the class, but you can use them during the breaks.

Class Schedule:

Week 1: Introduction to Research

Syllabus & Introductions.

Chapter 1.

The Research & Publication Process.

Week 2: Thinking Like a Researcher

Chapter 2.

A Tale of Two Papers:

- Fichman, R.G. and Kemerer, C.F., "The Illusory Diffusion of Innovation: An Examination of Assimilation Gaps," *Information Systems Research* (10:3), September 1999, pp. 255-275.
- Williams, L.; Kessler, R.R.; Cunningham, W.; and Jeffries, R., "Strengthening the Case for Pair Programming," *IEEE Software*, July/August 2000, pp. 19-25.

Critical Thinking:

- Dialog on Leadership, "Awareness is the First Critical Thing," A Conversation with Wanda Orlikowski, 1999.

Research Ethics:

- AIS Code of Conduct: <http://home.aisnet.org/displaycommon.cfm?an=1&subarticlenbr=13>
- IRB Process: http://www.research.usf.edu/cs/irb_forms.htm

Week 3: Literature Review & Meta-Analysis

Lit review:

- Fichman, R.G., "Information Technology Diffusion: A Review of Empirical Research," *Proceedings of the Thirteenth International Conference on Information Systems*, Dallas, 1992, 195-206.
- Alavi, M. and Leidner, D.E., "Knowledge Management and Knowledge Management Systems," *MIS Quarterly* (25:2), March 2001, pp. 107-136.

Meta-analysis:

- King, W.R. and He, J., "Understanding the Role and Methods of Meta-Analysis in IS Research," *Communications of the AIS* (16), 2005, pp. 665-686.
- Henard, D.H. and Szymanski, D.M., "Why Some New Products are More Successful Than Others," *Journal of Marketing Research* (38), August 2001, pp. 362-375.

Due: Research Proposal: Research Problem and its Significance.

Week 4: Philosophy of Science

Chapter 3.

Paradigms in scientific inquiry:

- Kuhn, T., *The Structure of Scientific Revolutions*, University of Chicago Press, 1996 (entire book).

[See [Dr. Pajares' notes](#) ONLY if you completely lost]

Social science paradigms:

- Krugman, P., "How Did Economists Get It So Wrong," *New York Times*, Sept 6, 2009.
- Gioia, D.A. and Pitre, E., "Multiparadigm Perspectives on Theory Building," *Academy of Management Review* (15:4), 1990, pp. 584-602.

Week 5: Theories in Organizational Research*Chapter 4.**Why theory:*

- Steinfield, C.W. and Fulk, J., "The Theory Imperative," in Organizations and Communications Technology, Janet Fulk and Charles W. Steinfield (eds.), Sage Publications, Newbury Park, CA, 1990.

Agency theory:

- Eisenhardt, K.M., "Agency Theory: An Assessment and Review," Academy of Management Review (14:1), 1989, pp. 57-74.

Transaction cost theory:

- Williamson, O.E., "The Economics of Organization: The Transaction Cost Approach," American Journal of Sociology (87:3), 1981, pp. 548-577.

Week 6: Organizational Theories - Continued*Resource-based theory:*

- Barney, J.B., "Firm Resources and Sustained Competitive Advantage," Journal of Management (17:1), 1991, pp. 99-120.
- Priem, R.L. and Butler, J.E., "Is the Resource-Based 'View' a Useful Perspective for Strategic Management Research?," Academy of Management Review (26:1), 2001, pp. 22-40.

Dynamic capability theory:

- Teece, D.J.; Pisano, G.; and Shuen, A., "Dynamic Capabilities and Strategic Management," Strategic Management Journal (18:7), 1997, 509-533.

Due: Research Proposal: Literature Review (with prior sections, modified if needed).

Week 7: Measurement and Validity*Chapters 6 and 7.**Scale validity and unidimensionality:*

- Straub, D.W., "Validating Instruments in MIS Research," MIS Quarterly (13:2), June 1989, pp. 146-169.
- MacKenzie, S. B., Podsakoff, P. M., and Podsakoff, N. P., "Construct Measurement and Validation Procedures in MIS and Behavioral Research: Integrating New and Existing Techniques," MIS Quarterly (35:2), 2011, pp. 293-334.

Due: Research Proposal: Theory & Hypotheses (with prior sections, modified as needed).

Week 8: Survey Research*Chapters 5, 8, and 9.**Field survey exemplar:*

- Tsai, W., "Knowledge Transfer in Intraorganizational Networks: Effects of Network Position and Absorptive Capacity on Business Unit Innovation and Performance," Academy of Management Review, 2001.

Biases in survey research:

- Malhotra, N. K, Kim, S. S., and Patil, A., "Common Method Variance in IS Research: A Comparison of Alternative Approaches and a Reanalysis of Past Research," Management Science (52:12), 2006, pp. 1865-1883.

Week 9: Experimental & Quasi-Experimental Research

Chapter 10.

Field experiment exemplar:

- Hunton, J.E. and McEwen, R.A., "An Assessment of the Relation Between Analysts' Earnings Forecast Accuracy, Motivational Incentives, and Cognitive Information Search Strategy," *The Accounting Review* (72:4), October 1997, pp. 497-515.

Problems with experimental research:

- Jarvenpaa, S. L.; Dickson, G. W.; and DeSanctis, G., "Methodological Issues in Experimental IS Research: Experiences and Recommendations," *MIS Quarterly*, June 1985, pp. 141-156.

Week 10: Reviewing Research

Writing paper reviews:

- Lee, A.S., "Reviewing a Manuscript for Publication," *Journal of Operations Management* (13:1), July 1995, pp.87-92.
- Agarwal, R; Echambadi, R; Franco, A.M.; and Sarkar, M.B., "Reap Rewards: Maximizing Benefits from Reviewer Comments," *Academy of Management Journal* (49:2), 2006, pp. 191-196.

Why papers are rejected:

- Daft, R.L., "Why I Recommended that Your Manuscript be Rejected and What You Can Do About It," in L.L. Cummings & P.J. Frost (eds.), *Publishing in the Organizational Sciences*, 2nd ed., 1995, pp. 164-182.

Formal review:

- Write a formal review of the following paper based on the guidelines prescribed in the above articles. Use this [review template](#) to structure your review: Anonymous, Paper submitted to *MIS Quarterly*, 2004. Actual comments from three reviewers, AE, and SE will be e-mailed to you the day before (after you complete your review).

Paper review:

- One "[mystery paper](#)" for you to review and submit on Week 13. This is part of your course grade, hence please plan to spend adequate time and effort on this review.

Week 11: Case Research

Chapter 11.

Conducting case research:

- Benbasat, I.; Goldstein, D.K.; and Mead, M., "The Case Research Strategy in Studies of Information Systems," *MIS Quarterly*, September 1987, pp. 369-386.

Case research exemplars:

- Beaudry, A. and Pinsonneault, A., "Understanding User Responses to Information Technology: A Coping Model of User Adaptation," *MIS Quarterly* (29:3), September 2005, pp. 493-524.
- Eisenhardt, K.M., "Making Fast Strategic Decisions In High-Velocity Environments," *Academy of Management Journal* (32:3), 1989, pp. 543-577.

Positivist versus Interpretive Analysis:

- Trauth, E.M. and Jessup, L.M., "Understanding Computer-Mediated Discussions: Positivist and Interpretive Analyses of Group Support System Use," *MIS Quarterly* (24:1), March 2000, pp. 43-79.

Due: Research Proposal: Research Methods (with all prior sections).

Week 12: Interpretive Research

Chapters 12 and 13.

Demo: Content Analysis using Grounded Theory.

Qualitative research:

- Shah, S.K. and Corley, K.G., "Building Better Theory by Bridging the Quantitative-Qualitative Divide," *Journal of Management Studies* (48:3), December 2006, pp. 1821-1835.

Action research:

- Kohli, R. and Kettinger, W., "Informing the Clan: Controlling Physician Costs and Outcomes," *MIS Quarterly* (28:3), September 2004, pp. 1-32.

Ethnography:

- Barley, S.R., "Technicians in the Workplace: Ethnographic Evidence for Bringing Work into Organization Studies," *Administrative Science Quarterly* (41), 1996, pp. 404-411.

Week 13: Miscellaneous Methods

Chapter 16.

Demo:

- Statistical analysis using SPSS.

Secondary data analysis:

- Chaney, P.K. and Philipich, K.L., "Shredded Reputation: The Cost of Audit Failure," *Journal of Accounting Research* (40:4), September 2002, pp. 1221-1245.

Analytic Modeling:

- Bayus, B.L., Jain, S., and Rao, A.G., "Truth or Consequences: An Analysis of Vaporware and New Product Announcements," *Journal of Marketing Research* (38), February 2001, pp. 3-13.

Due: Article Review.

Week 14: Finals

Comprehensive in-class 3-hour open-book exam. See [Sample Exam Questions](#).

Week 15: Student Presentations

20-minute presentation plus 5-minute Q & A.

Due: Final Research Proposal.

About the Book

This book is designed to introduce doctoral and graduate students to the process of conducting scientific research in the social sciences, business, education, public health, and related disciplines. It is a one-stop, comprehensive, and compact source for foundational concepts in behavioral research, and can serve as a stand-alone text or as a supplement to research readings in any doctoral seminar or research methods class. This book is currently being used as a research text at universities on six continents and will shortly be available in nine different languages. To receive updates on this book, including the translated versions, please follow the author on Facebook or Twitter @Anol_B.

About the Author

Anol Bhattacharjee is a professor of information systems and Citigroup/Hidden River Fellow at the University of South Florida, USA. He is one of the top ten information systems researchers in the world, ranked 8th based on research published in the top two journals in the discipline, MIS Quarterly and Information Systems Research, over the last decade (2001-2010). In a research career spanning 15 years, Dr. Bhattacharjee has published over 50 refereed journal papers and two books that have received over 4000 citations on Google Scholar. He also served on the editorial board of MIS Quarterly for four years and is frequently invited to present his research or build new research programs at universities all over the world. More information about Dr. Bhattacharjee can be obtained from his webpage at <http://ab2020.weebly.com>.

