

EXPERIMENTAL AND QUASI-EXPERIMENTAL DESIGNS FOR GENERALIZED CAUSAL INFERENCE

William R. Shadish

THE UNIVERSITY OF MEMPHIS

Thomas D. Cook

NORTHWESTERN UNIVERSITY

Donald T. Campbell

HOUGHTON MIFFLIN COMPANY

Boston New York

Contents

Preface	xv
1. EXPERIMENTS AND GENERALIZED CAUSAL INFERENCE	1
Experiments and Causation	3
Defining Cause, Effect, and Causal Relationships	3
Causation, Correlation, and Confounds	7
Manipulable and Nonmanipulable Causes	7
Causal Description and Causal Explanation	9
Modern Descriptions of Experiments	12
Randomized Experiment	13
Quasi-Experiment	13
Natural Experiment	17
Nonexperimental Designs	18
Experiments and the Generalization of Causal Connections	18
Most Experiments Are Highly Local But Have General Aspirations	18
Construct Validity: Causal Generalization as Representation	20
External Validity: Causal Generalization as Extrapolation	21
Approaches to Making Causal Generalizations	22
Experiments and Metascience	26
The Kuhnian Critique	27
Modern Social Psychological Critiques	28
Science and Trust	28
Implications for Experiments	29
A World Without Experiments or Causes?	31

2. STATISTICAL CONCLUSION VALIDITY AND INTERNAL VALIDITY	33
Validity	34
A Validity Typology	37
Threats to Validity	39
Statistical Conclusion Validity	42
Reporting Results of Statistical Tests of Covariation	42
Threats to Statistical Conclusion Validity	45
The Problem of Accepting the Null Hypothesis	52
Internal Validity	53
Threats to Internal Validity	54
Estimating Internal Validity in Randomized Experiments and Quasi-Experiments	61
The Relationship Between Internal Validity and Statistical Conclusion Validity	63
3. CONSTRUCT VALIDITY AND EXTERNAL VALIDITY	64
Construct Validity	64
Why Construct Inferences Are a Problem	66
Assessment of Sampling Particulars	69
Threats to Construct Validity	72
Construct Validity, Preexperimental Tailoring, and Postexperimental Specification	81
External Validity	83
Threats to External Validity	86
Constancy of Effect Size Versus Constancy of Causal Direction	90
Random Sampling and External Validity	91
Purposive Sampling and External Validity	92
More on Relationships, Tradeoffs, and Priorities	93
The Relationship Between Construct Validity and External Validity	93
The Relationship Between Internal Validity and Construct Validity	95
Tradeoffs and Priorities	96
Summary	102

4. QUASI-EXPERIMENTAL DESIGNS THAT EITHER LACK A CONTROL GROUP OR LACK PRETEST OBSERVATIONS ON THE OUTCOME	103
The Logic of Quasi-Experimentation in Brief	104
Designs Without Control Groups	106
The One-Group Posttest-Only Design	106
The One-Group Pretest-Posttest Design	108
The Removed-Treatment Design	111
The Repeated-Treatment Design	113
Designs That Use a Control Group But No Pretest	115
Posttest-Only Design With Nonequivalent Groups	115
Improving Designs Without Control Groups by Constructing Contrasts Other Than With Independent Control Groups	125
The Case-Control Design	128
Conclusion	134
5. QUASI-EXPERIMENTAL DESIGNS THAT USE BOTH CONTROL GROUPS AND PRETESTS	135
Designs That Use Both Control Groups and Pretests	136
The Untreated Control Group Design With Dependent Pretest and Posttest Samples	136
Matching Through Cohort Controls	148
Designs That Combine Many Design Elements	153
Untreated Matched Controls With Multiple Pretests and Posttests, Nonequivalent Dependent Variables, and Removed and Repeated Treatments	153
Combining Switching Replications With a Nonequivalent Control Group Design	154
An Untreated Control Group With a Double Pretest and Both Independent and Dependent Samples	154
The Elements of Design	156
Assignment	156
Measurement	158
Comparison Groups	159
Treatment	160
Design Elements and Ideal Quasi-Experimentation	160

Conclusion	161
Appendix 5.1: Important Developments in Analyzing Data From Designs With Nonequivalent Groups	161
Propensity Scores and Hidden Bias	162
Selection Bias Modeling	166
Latent Variable Structural Equation Modeling	169
6. QUASI-EXPERIMENTS: INTERRUPTED TIME-SERIES DESIGNS	171
What Is a Time Series?	172
Describing Types of Effects	172
Brief Comments on Analysis	174
Simple Interrupted Time Series	175
A Change in Intercept	175
A Change in Slope	176
Weak and Delayed Effects	178
The Usual Threats to Validity	179
Adding Other Design Features to the Basic Interrupted Time Series	181
Adding a Nonequivalent No-Treatment Control Group Time Series	182
Adding Nonequivalent Dependent Variables	184
Removing the Treatment at a Known Time	188
Adding Multiple Replications	190
Adding Switching Replications	192
Some Frequent Problems with Interrupted Time-Series Designs	195
Gradual Rather Than Abrupt Interventions	196
Delayed Causation	197
Short Time Series	198
Limitations of Much Archival Data	203
A Comment on Concomitant Time Series	205
Conclusion	206
7. REGRESSION DISCONTINUITY DESIGNS	207
The Basics of Regression Discontinuity	208
The Basic Structure	208

Examples of Regression Discontinuity Designs	212
Structural Requirements of the Design	216
Variations on the Basic Design	219
Theory of the Regression Discontinuity Design	220
Regression Discontinuities as Treatment Effects in the Randomized Experiment	221
Regression Discontinuity as a Complete Model of the Selection Process	224
Adherence to the Cutoff	227
Overrides of the Cutoff	227
Crossovers and Attrition	228
Fuzzy Regression Discontinuity	229
Threats to Validity	229
Regression Discontinuity and the Interrupted Time Series	229
Statistical Conclusion Validity and Misspecification of Functional Form	230
Internal Validity	237
Combining Regression Discontinuity and Randomized Experiments	238
Combining Regression Discontinuity and Quasi-Experiments	241
Regression Discontinuity—Experiment or Quasi-Experiment?	242
Appendix 7.1: The Logic of Statistical Proofs about Regression Discontinuity	243
 8. RANDOMIZED EXPERIMENTS: RATIONALE, DESIGNS, AND CONDITIONS CONDUCTIVE TO DOING THEM	 246
The Theory of Random Assignment	247
What Is Random Assignment?	248
Why Randomization Works	248
Random Assignment and Units of Randomization	253
The Limited Reach of Random Assignment	256
Some Designs Used with Random Assignment	257
The Basic Design	257
The Pretest-Posttest Control Group Design	261
Alternative-Treatments Design with Pretest	261

Multiple Treatments and Controls with Pretest	262
Factorial Designs	263
Longitudinal Designs	266
Crossover Designs	268
Conditions Most Conducive to Random Assignment	269
When Demand Outstrips Supply	269
When an Innovation Cannot Be Delivered to All Units at Once	270
When Experimental Units Can Be Temporally Isolated: The Equivalent-Time-Samples Design	270
When Experimental Units Are Spatially Separated or Interunit Communication Is Low	271
When Change Is Mandated and Solutions Are Acknowledged to Be Unknown	272
When a Tie Can Be Broken or Ambiguity About Need Can Be Resolved	273
When Some Persons Express No Preference Among Alternatives	273
When You Can Create Your Own Organization	274
When You Have Control over Experimental Units	274
When Lotteries Are Expected	275
When Random Assignment Is Not Feasible or Desirable	276
Discussion	277
9. PRACTICAL PROBLEMS 1: ETHICS, PARTICIPANT RECRUITMENT, AND RANDOM ASSIGNMENT	279
Ethical and Legal Issues with Experiments	280
The Ethics of Experimentation	281
Withholding a Potentially Effective Treatment	283
The Ethics of Random Assignment	286
Discontinuing Experiments for Ethical Reasons	289
Legal Problems in Experiments	290
Recruiting Participants to Be in the Experiment	292
Improving the Random Assignment Process	294
Methods of Randomization	294
What to Do If Pretest Means Differ	303
Matching and Stratifying	304
Matching and Analysis of Covariance	306
The Human Side of Random Assignment	307
Conclusion	311

Appendix 9.1: Random Assignment by Computer	311
SPSS and SAS	311
World Wide Web	313
Excel	313
10. PRACTICAL PROBLEMS 2: TREATMENT IMPLEMENTATION AND ATTRITION	314
Problems Related to Treatment Implementation	315
Inducing and Measuring Implementation	315
Analyses Taking Implementation into Account	320
Post-Assignment Attrition	323
Defining the Attrition Problem	323
Preventing Attrition	324
Analyses of Attrition	334
Discussion	340
11. GENERALIZED CAUSAL INFERENCE: A GROUNDED THEORY	341
The Received View of Generalized Causal Inference:	
Formal Sampling	342
Formal Sampling of Causes and Effects	344
Formal Sampling of Persons and Settings	346
Summary	348
A Grounded Theory of Generalized Causal Inference	348
Exemplars of How Scientists Make Generalizations	349
Five Principles of Generalized Causal Inferences	353
The Use of Purposive Sampling Strategies	354
Applying the Five Principles to Construct and External Validity	356
Should Experimenters Apply These Principles to All Studies?	371
Prospective and Retrospective Uses of These Principles	372
Discussion	373
12. GENERALIZED CAUSAL INFERENCE: METHODS FOR SINGLE STUDIES	374
Purposive Sampling and Generalized Causal Inference	374
Purposive Sampling of Typical Instances	375
Purposive Sampling of Heterogeneous Instances	376

Purposive Sampling and the First Four Principles	378
Statistical Methods for Generalizing from Purposive Samples	386
Methods for Studying Causal Explanation	389
Qualitative Methods	389
Statistical Models of Causal Explanation	392
Experiments That Manipulate Explanations	414
Conclusion	415
13. GENERALIZED CAUSAL INFERENCE: METHODS FOR MULTIPLE STUDIES	417
Generalizing from Single Versus Multiple Studies	418
Multistudy Programs of Research	418
Phased Models of Increasingly Generalizable Studies	419
Directed Programs of Experiments	420
Narrative Reviews of Existing Research	421
Narrative Reviews of Experiments	422
Narrative Reviews Combining Experimental and Nonexperimental Research	423
Problems with Narrative Reviews	423
Quantitative Reviews of Existing Research	425
The Basics of Meta-Analysis	426
Meta-Analysis and the Five Principles of Generalized Causal Inference	435
Discussion of Meta-Analysis	445
Appendix 13.1: Threats to the Validity of Meta-Analyses	446
Threats to Inferences About the Existence of a Relationship Between Treatment and Outcome	447
Threats to Inferences About the Causal Relationship Between Treatment and Outcome	450
Threats to Inferences About the Constructs Represented in Meta-Analyses	453
Threats to Inferences About External Validity in Meta-Analyses	454
14. A CRITICAL ASSESSMENT OF OUR ASSUMPTIONS	456
Causation and Experimentation	453
Causal Arrows and Pretzels	453
Epistemological Criticisms of Experiments	459
Neglected Ancillary Questions	461

Validity	462
Objections to Internal Validity	462
Objections Concerning the Discrimination Between Construct Validity and External Validity	466
Objections About the Completeness of the Typology	473
Objections Concerning the Nature of Validity	475
Quasi-Experimentation	484
Criteria for Ruling Out Threats: The Centrality of Fuzzy Plausibility	484
Pattern Matching as a Problematic Criterion	485
The Excuse Not to Do a Randomized Experiment	486
Randomized Experiments	488
Experiments Cannot Be Successfully Implemented	488
Experimentation Needs Strong Theory and Standardized Treatment Implementation	489
Experiments Entail Tradeoffs Not Worth Making	490
Experiments Assume an Invalid Model of Research Utilization	493
The Conditions of Experimentation Differ from the Conditions of Policy Implementation	495
Imposing Treatments Is Fundamentally Flawed Compared with Encouraging the Growth of Local Solutions to Problems	497
Causal Generalization: An Overly Complicated Theory?	498
Nonexperimental Alternatives	499
Intensive Qualitative Case Studies	500
Theory-Based Evaluations	501
Weaker Quasi-Experiments	502
Statistical Controls	503
Conclusion	504
Glossary	505
References	514
Name Index	593
Subject Index	609