

RealWorld Evaluation

Working Under Budget, Time, Data, and Political Constraints

Third Edition

Michael Bamberger

Independent Consultant, Former World Bank Senior Sociologist

Linda Mabry

Washington State University

Foreword by Jim Rugh



Los Angeles | London | New Delhi
Singapore | Washington DC | Melbourne

BRIEF CONTENTS

List of Boxes, Figures, and Tables	xxiv
List of Appendices	xxxix
Foreword by Jim Rugh	xxxiii
Preface	xxxv
Acknowledgments	xliv
About the Authors	xliv

PART I • THE SEVEN STEPS OF THE REALWORLD EVALUATION APPROACH

Chapter 1 • Overview: RealWorld Evaluation and the Contexts in Which It Is Used	2
Chapter 2 • First Clarify the Purpose: Scoping the Evaluation	19
Chapter 3 • Not Enough Money: Addressing Budget Constraints	39
Chapter 4 • Not Enough Time: Addressing Scheduling and Other Time Constraints	55
Chapter 5 • Critical Information Is Missing or Difficult to Collect: Addressing Data Constraints	72
Chapter 6 • Political Constraints	94
Chapter 7 • Strengthening the Evaluation Design and the Validity of the Conclusions	105
Chapter 8 • Making It Useful: Helping Clients and Other Stakeholders Utilize the Evaluation	127

PART II • A REVIEW OF EVALUATION METHODS AND APPROACHES AND THEIR APPLICATION IN REALWORLD EVALUATION: FOR THOSE WHO WOULD LIKE TO DIG DEEPER

Chapter 9 • Standards and Ethics	144
Chapter 10 • Theory-Based Evaluation and Theory of Change	153
Chapter 11 • Evaluation Designs: The RWE Strategy for Selecting the Appropriate Evaluation Design to Respond to the Purpose and Context of Each Evaluation	187

Chapter 12 • Quantitative Evaluation Methods	218
Chapter 13 • Qualitative Evaluation Methods	243
Chapter 14 • Mixed-Method Evaluation	262
Chapter 15 • Sampling Strategies for RealWorld Evaluation	289
Chapter 16 • Evaluating Complex Projects, Programs, and Policies	315
Chapter 17 • Gender Evaluation: Integrating Gender Analysis Into Evaluations	353
Chapter 18 • Evaluation in the Age of Big Data	382
 PART III • MANAGING EVALUATIONS	
Chapter 19 • Managing Evaluations	430
Chapter 20 • The Road Ahead	459
 Glossary of Terms and Acronyms	474
References	487
Author Index	504
Subject Index	510

DETAILED CONTENTS

List of Boxes, Figures, and Tables	xxiv
List of Appendices	xxxi
Foreword by Jim Rugh	xxxiii
Preface	xxxv
Acknowledgments	xliv
About the Authors	xliv

PART I • THE SEVEN STEPS OF THE REALWORLD EVALUATION APPROACH

Chapter 1 • Overview: RealWorld Evaluation and the Contexts in Which It Is Used	2
1. Welcome to RealWorld Evaluation	2
2. The RealWorld Evaluation Context	8
3. The Four Types of Constraints Addressed by the RealWorld Approach	9
3.1 Budget and Other Resource Constraints	9
3.2 Time Constraints	11
3.3 Data Constraints	11
3.4 Political and Organizational Influences and Constraints	12
4. Additional Organizational and Administrative Challenges	13
5. The RealWorld Approach to Evaluation Challenges	14
6. Who Uses RealWorld Evaluation, for What Purposes, and When?	16
Summary	18
Further Reading	18
Chapter 2 • First Clarify the Purpose: Scoping the Evaluation	19
1. Stakeholder Expectations of Impact Evaluations	20
2. Understanding Information Needs	22
3. Developing the Program Theory Model	25
3.1 Theory-Based Evaluation (TBE) as a Management Tool	30
4. Identifying the Constraints to Be Addressed by RWE and Determining the Appropriate Evaluation Design	30
5. Developing Designs Suitable for RealWorld Evaluation Conditions	31
5.1 How the Availability of Data Affects the Choice of Evaluation Design	32
5.2 Developing the Terms of Reference (Statement of Work) for the Evaluation	36
Summary	37
Further Reading	37

Chapter 3 • Not Enough Money: Addressing Budget Constraints	39
1. Simplifying the Evaluation Design	39
1.1 Simplifying the Design for Quantitative Evaluations	42
1.2 Simplifying the Design for Qualitative Evaluations	44
2. Clarifying Client Information Needs	46
3. Using Existing Data	47
4. Reducing Costs by Reducing Sample Size	47
4.1 Adjusting the Sample Size Based on Client Information Needs and the Kinds of Decisions to Which the Evaluation Will Contribute	47
4.2 Factors Affecting Sample Size for Quantitative Evaluations	48
<i>Effect of the Level of Disaggregation on the Required Sample Size</i>	49
4.3 Factors Affecting the Size of Qualitative Samples	50
4.4 Factors Affecting the Size of Mixed-Method Samples	50
5. Reducing Costs of Data Collection and Analysis	50
6. Assessing the Feasibility and Utility of Using New Information Technology (NIT) to Reduce the Costs of Data Collection	52
7. Threats to Validity of Budget Constraints	53
Summary	53
Further Reading	54
 Chapter 4 • Not Enough Time: Addressing Scheduling and Other Time Constraints	 55
1. Similarities and Differences Between Time and Budget Constraints	55
2. Simplifying the Evaluation Design	59
3. Clarifying Client Information Needs and Deadlines	60
4. Using Existing Documentary Data	61
5. Reducing Sample Size	61
6. Rapid Data-Collection Methods	61
7. Reducing Time Pressure on Outside Consultants	67
8. Hiring More Resource People	68
9. Building Outcome Indicators Into Project Records	68
10. New Information Technology for Data Collection and Analysis	69
11. Common Threats to Adequacy and Validity Relating to Time Constraints	70
Summary	70
Further Reading	71
 Chapter 5 • Critical Information Is Missing or Difficult to Collect: Addressing Data Constraints	 72
1. Data Issues Facing RealWorld Evaluators	72
2. Reconstructing Baseline Data	76
2.1 Strategies for Reconstructing Baseline Data	76
<i>When Should Baseline Data Be Collected?</i>	76
<i>Using Administrative Data From the Project</i>	77
<i>Cautionary Tales—and Healthy Skepticism</i>	79
<i>Using Other Sources of Secondary Data</i>	80

<i>Conducting Retrospective Surveys</i>	81
<i>Working With Key Informants</i>	83
<i>Using Participatory Evaluation Methods</i>	83
<i>Using Geographical Information Systems (GIS) and Satellite Images to Reconstruct Baseline Data</i>	84
3. Special Issues Reconstructing Baseline Data for Project Populations and Comparison Groups	85
3.1 Challenges Collecting Baseline Data on the Project Group	85
3.2 Challenges Collecting Baseline Data on a Comparison Group	86
3.3 The Challenge of Omitted Variables ("Unobservables")	86
4. Collecting Data on Sensitive Topics or From Difficult-to-Reach Groups	88
4.1 Addressing Sensitive Topics	88
4.2 Studying Difficult-to-Reach Groups	88
5. Common Threats to Adequacy and Validity of an Evaluation Relating to Data Constraints	92
Summary	92
Further Reading	93
Chapter 6 • Political Constraints	94
1. Values, Ethics, and Politics	94
2. Societal Politics and Evaluation	95
3. Stakeholder Politics	97
4. Professional Politics	98
4.1 Teaming	98
5. Political Issues in the Design Phase	99
5.1 Hidden Agendas and Pseudo-Evaluation	99
5.2 Stakeholder Differences	100
6. Political Issues in the Conduct of an Evaluation	100
6.1 Shifting Evaluator Roles	100
6.2 Data Access	101
7. Political Issues in Evaluation Reporting and Use	101
7.1 Bias in Reporting	102
7.2 Use and Misuse of Findings	102
8. Advocacy	102
8.1 Activism	103
8.2 Strategizing for Use	103
Summary	104
Further Reading	104
Chapter 7 • Strengthening the Evaluation Design and the Validity of the Conclusions	105
1. Validity in Evaluation	105
2. Factors Affecting Adequacy and Validity	107
3. A Framework for Assessing the Validity and Adequacy of QUANT, QUAL, and Mixed-Method Designs	108
3.1 The Categories of Validity (Adequacy, Trustworthiness)	109

4. Assessing and Addressing Threats to Validity for Quantitative Impact Evaluations	112
4.1 A Threats-to-Validity Worksheet for QUANT Evaluations	112
4.2 Strengthening Validity in Quantitative Evaluations by Strengthening the Evaluation Design	114
<i>Random Sampling</i>	114
<i>Triangulation</i>	115
<i>Selection of Statistical Procedures</i>	116
<i>Peer Review and Meta-Evaluation</i>	116
4.3 Taking Corrective Actions When Threats to Validity Have Been Identified	116
5. Assessing Adequacy and Validity for Qualitative Impact Evaluations	118
5.1 Strengthening Validity in Qualitative Evaluations by Strengthening the Evaluation Design	120
<i>Purposeful (Purposive) Sampling</i>	120
<i>Triangulation</i>	120
<i>Validation</i>	121
<i>Meta-Evaluation and Peer Review</i>	121
5.2 Addressing Threats to Validity in Qualitative Evaluations	121
<i>Collecting Data Across the Full Range of Appropriate Settings, Times, and Respondents</i>	121
<i>Inappropriate Subject Selection</i>	122
<i>Insufficient Language or Cultural Skills to Ensure Sensitivity to Informants</i>	122
<i>Insufficient Opportunity for Ongoing Analysis by the Team</i>	122
<i>Minimizing Observer Effects</i>	122
6. Assessing Validity for Mixed-Method (MM) Evaluations	122
7. Using the Threats-to-Validity Worksheets	123
7.1 Points During the RWE Cycle at Which Corrective Measures Can Be Taken	124
<i>Strengthening the Evaluation Design</i>	125
Summary	125
Further Reading	126

Chapter 8 • Making It Useful: Helping Clients and Other Stakeholders Utilize the Evaluation	127
1. What Do We Mean by Influential Evaluations and Useful Evaluations?	127
2. The Underutilization of Evaluation Studies	130
2.1 Why Are Evaluation Findings Underutilized?	131
2.2 The Challenges of Utilization for RWE	131
3. Strategies for Promoting the Utilization of Evaluation Findings and Recommendations	132
3.1 The Importance of the Scoping Phase	133
<i>Understand the Client's Information Needs</i>	133
<i>Understand the Dynamics of the Decision-Making Process and the Timing of the Different Steps</i>	135
<i>Define the Program Theory on Which the Program Is Based in Close Collaboration With Key Stakeholders</i>	135
<i>Identify Budget, Time, and Data Constraints and Prioritize Their Importance and the Client's Flexibility to Adjust Budget or Time If Required to Improve the Quality of the Evaluation</i>	136

<i>Understand the Political Context</i>	136
<i>Prepare a Set of RWE Design Options to Address the Constraints and to Strategize</i>	
<i>With the Client to Assess Which Option Is Most Acceptable</i>	136
3.2 Formative Evaluation Strategies	136
3.3 Communication With Clients Throughout the Evaluation	138
3.4 Evaluation Capacity Building	138
3.5 Strategies for Overcoming Political and Bureaucratic Challenges	138
3.6 Communicating Findings	139
3.7 Developing a Follow-Up Action Plan	141
Summary	141
Further Reading	141

PART II • A REVIEW OF EVALUATION METHODS AND APPROACHES AND THEIR APPLICATION IN REALWORLD EVALUATION: FOR THOSE WHO WOULD LIKE TO DIG DEEPER

Chapter 9 • Standards and Ethics	144
1. Standards of Competence	144
2. Professional Standards	144
2.1 Evaluation Standards	145
2.2 Guiding Principles	146
3. Ethical Codes of Conduct	146
3.1 International Standards of Ethics	147
3.2 National Standards of Ethics	147
3.3 Ethical Frameworks in Evaluation	148
4. Issues	149
4.1 Advocacy	149
4.2 Accountability	149
4.3 International Regulation	150
Summary	151
Further Reading	151
Chapter 10 • Theory-Based Evaluation and Theory of Change	153
1. Theory-Based Evaluation (TBE) and Theory of Change (TOC)	153
1.1 Program Theory and Theory-Based Evaluation	153
<i>Stage 1: Articulation of the Program Theory Model</i>	154
<i>Stage 2: The Results Framework</i>	154
<i>Stage 3: The Logical Framework</i>	154
<i>Stages 4a and 4b: Impact and Implementation Models</i>	156
<i>Applications of Program Theory</i>	156
1.2 Theory of Change	156
<i>Benefits of the TOC for Program Evaluation</i>	158
2. Applications of Program Theory in Program Evaluation	160
2.1 The Increasing Use of Program Theory in Evaluation	160
2.2 Examples of TBE Tools and Techniques That Can Be Applied in Policy, Program, and Project Evaluations and at Different Stages of the Program Cycle	162
2.3 Realist Evaluation	164

3. Using TOC in Program Evaluation	165
3.1 Conceptualizing the Change Process	165
3.2 Identifying and Testing Hypotheses About the Processes of Implementation and Change	165
3.3 Identifying Unintended Outcomes	165
3.4 Addressing Complexity and Emergence	166
<i>Addressing Emergence</i>	167
3.5 Challenges Affecting the Use of TOCs	167
4. Designing a Theory of Change Evaluation Framework	168
4.1 The Different Ways That a Theory of Change Can Be Used	168
4.2 The Purpose of the Theory of Change	169
4.3 Representing the Theory of Change	170
<i>An Example of a Theory of Change</i>	172
4.4 Designing the TOC and the Sources of Data	172
5. Integrating a Theory of Change Into the Program Management, Monitoring, and Evaluation Cycle	175
5.1 Articulating the Program Theory	175
5.2 Process Tracing	177
5.3 Results-Based Reporting and Logical Frameworks	177
5.4 Program Impact and Implementation Models	179
6. Program Theory Evaluation and Causality	179
6.1 The Debate on the Value of Program Theory to Explain Causality	179
6.2 Using Mixed-Method Program Theory Designs to Explain Causality	181
Summary	184
Further Reading	184

Chapter 11 • Evaluation Designs: The RWE Strategy for Selecting the Appropriate Evaluation Design to Respond to the Purpose and Context of Each Evaluation	187
1. Different Approaches to the Classification of Evaluation Designs	187
2. Assessing Causality Attribution and Contribution	190
2.1 Attribution Analysis and Contribution Analysis	190
2.2 An Introduction to Contribution Analysis and Outcome Harvesting	190
3. The RWE Approach to the Selection of the Appropriate Impact Evaluation Design	192
3.1 Design Step 1	193
<i>Using the Evaluation Purpose and Context Checklist</i>	193
3.2 Design Step 2	197
<i>Analysis of the Evaluation Design Framework</i>	197
3.3 Design Step 3	201
<i>Identify a Short List of Potential Evaluation Designs</i>	201
3.4 Design Step 4	204
<i>Take Into Consideration the Preferred Methodological Approaches of Stakeholders and Evaluators</i>	204
3.5 Design Step 5	205
<i>Strategies for Strengthening the Basic Evaluation Designs</i>	205
3.6 Design Step 6	205
<i>Evaluability Analysis to Assess the Technical, Resource, and Political Feasibility of Each Design</i>	205

3.7 Design Step 7	205
<i>Preparation of Short List of Evaluation Design Options for Discussion With Clients and Other Stakeholders</i>	205
3.8 Design Step 8	207
<i>Agreement on the Final Evaluation Design</i>	207
4. Tools and Techniques for Strengthening the Basic Evaluation Designs	208
4.1 Basing the Evaluation on a Theory of Change and a Program Theory Model	208
4.2 Process Analysis	209
4.3 Incorporating Contextual Analysis	209
4.4 Complementing Quantitative Data Collection and Analysis With Mixed-Method Designs	210
4.5 Making Full Use of Available Secondary Data	210
4.6 Triangulation: Using Two or More Independent Estimates for Key Indicators and Using Data Sources and Analytical Methods to Explain Findings	210
5. Selecting the Best Design for RealWorld Evaluation Scenarios	210
5.1 Factors Affecting the Choice of the Appropriate Design for a Particular Evaluation	210
5.2 Challenges Facing the Use of Experimental and Other Statistical Designs in RealWorld Evaluation Contexts	213
5.3 Selecting the Appropriate Designs for RealWorld Evaluation Scenarios	213
Summary	215
Further Reading	216
Chapter 12 • Quantitative Evaluation Methods	218
1. Quantitative Evaluation Methodologies	218
1.1 The Importance of Program Theory in the Design and Analysis of QUANT Evaluations	218
1.2 Quantitative Sampling	219
2. Experimental and Quasi-Experimental Designs	219
2.1 Randomized Control Trials (RCTs)	219
2.2 Quasi-Experimental Designs (QEDs)	223
3. Strengths and Weaknesses of Quantitative Evaluation Methodologies	226
4. Applications of Quantitative Methodologies in Program Evaluation	226
4.1 Analysis of Population Characteristics	226
4.2 Hypothesis Testing and the Analysis of Causality	226
4.3 Cost-Benefit Analysis and the Economic Rate of Return (ERR)	228
4.4 Cost-Effectiveness Analysis	228
5. Quantitative Methods for Data Collection	229
5.1 Questionnaires	229
<i>Types of Questions Used in Quantitative Surveys</i>	230
5.2 Interviewing	230
5.3 Observation	230
<i>Observational Protocols</i>	230
<i>Unobtrusive Measures in Observation</i>	231
5.4 Focus Groups	231
5.5 Self-Reporting Methods	232
5.6 Knowledge and Achievement Tests	232

5.7 Anthropometric and Other Physiological Health Status Measures	232
5.8 Using Secondary Data	233
<i>Common Problems With Secondary Data for Evaluation Purposes</i>	234
6. The Management of Data Collection for Quantitative Studies	235
6.1 Survey Planning and Design	235
6.2 Implementation and Management of Data Collection	236
<i>RealWorld Constraints on the Management of Data Collection</i>	237
7. Data Analysis	239
Summary	240
Further Reading	240
Chapter 13 • Qualitative Evaluation Methods	243
1. Design	243
1.1 Flexibility	244
1.2 Contextuality	245
1.3 Sensitivity	245
2. Data Collection	245
2.1 Observation	246
<i>Collecting Observation Data</i>	246
<i>Structure in Observations</i>	246
<i>Participant Observations</i>	247
<i>Writing Up Observation Data</i>	250
2.2 Interview	250
<i>Collecting Interview Data</i>	250
<i>Structure in Interviews</i>	251
<i>Group Interviews</i>	251
<i>Hybrid Interview Methods</i>	252
<i>Writing Up Interview Data</i>	253
2.3 Analysis of Documents and Artifacts	253
2.4 Technology in Data Collection	254
2.5 Triangulation and Validation	254
<i>Triangulation</i>	255
<i>Validation</i>	255
3. Data Analysis	256
3.1 Inductive Analysis	256
<i>Constant-Comparative Method</i>	256
3.2 Thematic Analysis	256
<i>Electronic Data Analysis</i>	256
3.3 Holistic Analysis	257
<i>Data Complexity</i>	257
<i>Criteriality</i>	257
3.4 Validity	257
<i>Disciplining Subjective Judgment</i>	258
<i>Theoretical Triangulation</i>	258
<i>Peer Review and Meta-Evaluation</i>	258
3.5 Generalizability	259
4. Reporting	259
5. Real-World Constraints	260

Summary	260
Further Reading	261
Chapter 14 • Mixed-Method Evaluation	262
1. The Mixed-Method Approach	262
2. Rationale for Mixed-Method Approaches	263
2.1 Why Use Mixed Methods?	263
2.2 Areas Where Mixed Methods Can Potentially Strengthen Evaluations	267
<i>Improving Construct Validity and Data Quality</i>	267
<i>Evaluating Complex Programs</i>	268
<i>Strengthening Big Data–Based Evaluations</i>	268
<i>Identifying Unintended Outcomes (UOs) of Development Programs</i>	269
3. Approaches to the Use of Mixed Methods	269
3.1 Applying Mixed Methods When the Dominant Design Is Quantitative or Qualitative	273
3.2 Using Mixed Methods When Working Under Budget, Time, and Data Constraints	275
4. Mixed-Method Strategies	276
4.1 Sequential Mixed-Method Designs	276
4.2 Concurrent Designs	278
<i>Concurrent Triangulation Design</i>	279
<i>Concurrent Nested Design</i>	280
4.3 Using Mixed Methods at Different Stages of the Evaluation	281
5. Implementing a Mixed-Method Design	281
<i>Composition of the Research Team</i>	283
<i>Using Integrated Approaches at Different Stages of the Evaluation</i>	283
6. Using Mixed Methods to Tell a More Compelling Story of What a Program Has Achieved	284
7. Case Studies Illustrating the Use of Mixed Methods	286
Summary	286
Further Reading	287
Chapter 15 • Sampling Strategies for RealWorld Evaluation	289
1. The Importance of Sampling for RealWorld Evaluation	290
2. Purposive Sampling	291
2.1 Purposive Sampling Strategies	292
2.2 Purposive Sampling for Different Types of Qualitative Data Collection and Use	294
<i>Sampling for Data Collection</i>	294
<i>Sampling Data for Qualitative Reporting</i>	295
2.3 Considerations in Planning Purposive Sampling	295
3. Probability (Random) Sampling	296
3.1 Key Questions in Designing a Random Sample for Program Evaluation	296
3.2 Selection Procedures in Probability (Random) Sampling	299
3.3 Sample Design Decisions at Different Stages of the Survey	300
<i>Presampling Questions</i>	300
<i>Questions and Choices During the Sample Design Process</i>	301
<i>Postsampling Questions and Choices</i>	302

3.4 Sources of Error in Probabilistic Sample Design	303
<i>Nonsampling Bias</i>	303
<i>Sampling Bias</i>	304
4. Using Power Analysis and Effect Size for Estimating the Appropriate Sample Size for an Impact Evaluation	305
4.1 The Importance of Power Analysis for Determining Sample Size for Probability Sampling	305
4.2 Defining Effect Size	306
4.3 Type I and Type II Errors	306
4.4 Defining the Power of the Test	307
4.5 An Example Illustrating the Relationship Between the Power of the Test, the Effect Size, and the Required Sample Size	307
5. The Contribution of Meta-Analysis	309
6. Sampling Issues for Mixed-Method Evaluations	309
6.1 Model 1: Using Mixed Methods to Strengthen a Mainly Quantitative Evaluation Design	309
6.2 Model 2: Using a Mixed-Method Design to Strengthen a Qualitative Evaluation Design	311
6.3 Model 3: Using an Integrated Mixed-Method Design	311
7. Sampling Issues for RealWorld Evaluation	312
7.1 Lot Quality Acceptance Sampling (LQAS): An Example of a Sampling Strategy Designed to Be Economical and Simple to Administer and Interpret	312
Summary	313
Further Reading	314
Chapter 16 • Evaluating Complex Projects, Programs, and Policies	315
1. The Move Toward Complex, Country-Level Development Programming	315
2. Defining Complexity in Development Programs and Evaluations	317
2.1 The Dimensions of Complexity	317
<i>Dimension 1: The Nature of the Intervention</i>	317
<i>Dimension 2: Institutions and Stakeholders</i>	319
<i>Dimension 3: The Context (System) Within Which the Program Is Implemented</i>	319
<i>Dimension 4: Causality and Change</i>	320
<i>The Complexity of the Evaluation</i>	320
2.2 Simple Projects, Complicated Programs, and Complex Development Interventions	321
2.3 The Main Types of Complex Interventions	326
2.4 Assessing Levels of Complexity	327
<i>How to Use the Checklist</i>	330
2.5 Special Challenges for the Evaluation of Complex, Country-Level Programs	330
3. A Framework for the Evaluation of Complex Development Programs	331
3.1 Overview of the Complexity-Responsive Evaluation Framework	331
3.2 Drawing on Big Data Science	333
3.3 Step 1: Mapping the Dimensions of Complexity	333
3.4 Step 2: Choosing a Unit of Analysis for Unpacking Complex Interventions Into Evaluable Components	338

<i>Option 1: Implementation Components, Phases, and Themes</i>	338
<i>Option 2: Program Theories</i>	339
<i>Option 3: Cases</i>	339
<i>Option 4: Variables</i>	340
3.5 Step 3: Choosing an Evaluation Design for Evaluating Each	
Unpacked Component	340
<i>Attribution and the Challenge of Defining the Counterfactual for Complex Evaluations</i>	340
<i>Choosing the Best Evaluation Design for Evaluating the Unpacked Program Components</i>	340
<i>Operational Unpacking of Components of Complex Interventions</i>	342
<i>Program Theory and Theory Reconstruction</i>	343
<i>Case Studies and Rich Description</i>	344
<i>Variable-Based Approaches to Unpacking</i>	344
3.6 Step 4: Choosing an Approach to Reassembling the Various Parts Into a Big Picture	345
<i>Systems Modeling</i>	345
<i>Descriptive and Inferential Statistical Analysis</i>	346
<i>Comparative Case Study Approaches</i>	346
<i>Portfolio Analysis</i>	347
<i>Review and Synthesis Approaches</i>	347
<i>Rating Scales</i>	348
3.7 Step 5: Assessing Program Contribution to National and International	
Development Goals—Going Back to the Big Picture	350
Summary	351
Further Reading	352

Chapter 17 • Gender Evaluation: Integrating Gender Analysis	
Into Evaluations	353
1. Why a Gender Focus Is Critical	354
1.1 Why So Few Evaluations Have a Gender Focus	355
<i>Political Constraints</i>	355
<i>Skill and Resource Constraints</i>	355
<i>Methodological Constraints</i>	356
1.2 The Value of a Gender Evaluation	356
2. Gender Issues in Evaluations	357
<i>Gender Does Not Just Focus on Women</i>	357
<i>Intersectionality</i>	357
<i>Multisectoral Analysis</i>	357
<i>Complexity and a Longitudinal Perspective</i>	358
<i>Gender Approaches Are Normative</i>	358
<i>Bodily Integrity and Sexuality</i>	359
<i>Systems of Social Control and Gender</i>	359
<i>Defining Boundaries for the Evaluation</i>	360
2.1 How to “Gender” an Evaluation	361
<i>Using a Gender Analysis Framework</i>	361
<i>Using a Feminist Evaluation Approach</i>	363
2.2 A Continuum of Gendered Evaluation	363
<i>Level 1: Sex Disaggregation of a Set of Basic Indicators</i>	364
<i>Level 2: Analysis of Factors Affecting Women’s Participation in Development</i>	365
<i>Level 3: Analysis of Household, Community, and Social Dynamics</i>	365
<i>Level 4: Comprehensive Feminist Analysis for Empowerment</i>	366

3. Designing a Gender Evaluation	366
3.1 Design Criteria	366
<i>Evaluation Criteria—Adapting the OECD/DAC Framework</i>	366
3.2 Selecting Projects for a Gendered Evaluation	367
<i>Institutional Intervention Points</i>	368
<i>Portfolio Analysis and Meta-Analysis</i>	368
<i>Gender Flags and Checklists</i>	368
3.3 Defining the Depth and Scope	369
<i>Depth of the Analysis</i>	369
<i>Boundaries of the Evaluation</i>	369
<i>Time Horizons</i>	370
3.4 Defining Evaluation Questions	371
<i>The Importance of Broad-Based Stakeholder Consultations</i>	372
3.5 Evaluability Assessment	372
4. Gender Evaluations With Different Scopes	373
4.1 Complex Development Programs Evaluation	373
4.2 Single-Country Evaluations	373
4.3 Multicountry Evaluations	374
4.4 Country Gender M&E Framework	374
4.5 Gender in Sustainable Development Goals	374
5. The Tools of Gender Evaluation	376
5.1 Advanced Designs for Gender Evaluations	377
5.2 Attribution and Contribution Analysis	378
Summary	378
Further Reading	379

Chapter 18 • Evaluation in the Age of Big Data **382**

1. Introducing Big Data and Data Science	383
1.1 Increasing Application of Big Data in Personal and Public Life	383
1.2 Defining Big Data, Data Science, and New Information Technology (NIT)	384
<i>Big Data Is Embedded in New Information Technology</i>	384
<i>Defining Big Data</i>	387
1.3 The Data Continuum	392
1.4 The Position of Evaluation in the Big Data Ecosystem	393
<i>The Big Data Ecosystem</i>	393
<i>The Evaluation Ecosystem</i>	394
2. Increasing Application of Big Data in the Development Context	396
3. The Tools of Data Science	398
3.1 The Stages of the Data Analytics Cycle	398
<i>Step 1: Descriptive and Exploratory Analysis: Documenting What Is Happening, Often in Real Time</i>	399
<i>Step 2: Predictive Analysis: What Is Likely to Happen</i>	401
<i>Step 3: Detection: Tracking Who Is Likely to Succeed and Who Will Fail</i>	403
<i>Step 4: Prescription: Evaluating How Outcomes Were Achieved and Providing Recommendations on How to Improve Program Performance</i>	403
3.2 The Tools of Data Analytics	404
3.3 The Limitations of Data Science	408
<i>Being Aware of, and Addressing, Methodological Challenges</i>	408
<i>Political and Organizational Challenges</i>	408

<i>Privacy and Security</i>	409
<i>New Ethical Challenges</i>	409
<i>The Dark Side of Big Data</i>	410
4. Potential Applications of Data Science in Development Evaluation	411
4.1 Challenges Facing Current Evaluation Approaches	411
<i>Data Analysis Challenges</i>	412
4.2 How Data Science Can Strengthen Evaluation Practice	413
5. Building Bridges Between Data Science and Evaluation	417
5.1 The Potential Benefits of Convergence of Data Science and Evaluation	417
5.2 The Need for Caution When Assessing the Benefits of Big Data and Data Analytics for Development Evaluation	419
5.3 Challenges	420
5.4 Skills Required for Present and Future Evaluators	422
<i>Understanding the Big Debates Around the Evaluation of Future Development Programs</i>	422
<i>New Skills Required for Evaluation Offices and Evaluators and for Data Scientists</i>	423
<i>Skills Development for Evaluators</i>	423
<i>Evaluation Skills for Data Scientists</i>	424
5.5 Necessary Conditions for Convergence to Occur	424
<i>To What Extent Do These Conditions Exist in Both Industrial and Developing Countries?</i>	425
5.6 Possible Collaborative Bridge-Building Initiatives	425
Summary	425
Further Reading	426

PART III • MANAGING EVALUATIONS

Chapter 19 • Managing Evaluations 430

1. Organizational and Political Issues Affecting the Design, Implementation, and Use of Evaluations	430
2. Planning and Managing the Evaluation	431
2.1 Step 1: Preparing the Evaluation	431
<i>Step 1-A: Defining the Evaluation Framework or the Scope of Work (SoW)</i>	431
<i>Step 1-B: Linking the Evaluation to the Project Results Framework</i>	433
<i>Step 1-C: Involving Stakeholders</i>	433
<i>Step 1-D: Commissioning Diagnostic Studies</i>	433
<i>Step 1-E: Defining the Management Structure for the Evaluation</i>	435
2.2 Step 2: Recruiting the Evaluators	436
<i>Step 2-A: Recruiting the Internal Evaluation Team</i>	436
<i>Step 2-B: Different Ways to Contract External Evaluation Consultants</i>	436
<i>Step 2-C: Preparing the Request for Proposals (RFP)</i>	437
<i>Step 2-D: Preparing the Terms of Reference (ToR)</i>	438
<i>Step 2-E: Selecting the Consultants</i>	439
2.3 Step 3: Designing the Evaluation	439
<i>Step 3-A: Formulating Evaluation Questions</i>	439
<i>Step 3-B: Assessing the Evaluation Scenario</i>	441
<i>Step 3-C: Selecting the Appropriate Evaluation Design</i>	441
<i>Step 3-D: Commissioning an Evaluability Assessment</i>	441
<i>Step 3-E: Designing "Evaluation-Ready" Programs</i>	442

2.4 Step 4: Implementing the Evaluation	443
<i>Step 4-A: The Role of the Evaluation Department of a Funding Agency, Central Government Ministry, or Sector Agency</i>	443
<i>Step 4-B: The Inception Report</i>	444
<i>Step 4-C: Managing the Evaluation</i>	444
<i>Step 4-D: Working With Stakeholders</i>	445
<i>Step 4-E: Quality Assurance (QA)</i>	446
<i>Step 4-F: Management Challenges With Different Kinds of Evaluation</i>	447
2.5 Step 5: Reporting and Dissemination	449
<i>Step 5-A: Providing Feedback on the Draft Report</i>	449
<i>Step 5-B: Disseminating the Evaluation Report</i>	450
2.6 Step 6: Ensuring Implementation of the Recommendations	450
<i>Step 6-A: Coordinating the Management Response and Follow-Up</i>	450
<i>Step 6-B: Facilitating Dialogue With Partners</i>	451
3. Institutionalizing Impact Evaluation Systems at the Country and Sector Levels	451
3.1 Institutionalizing Impact Evaluation	451
3.2 Integrating IE Into Sector and/or National M&E and Other Data-Collection Systems	453
<i>Creating Demand for IE</i>	454
4. Evaluating Capacity Development	454
Summary	455
Further Reading	457
 Chapter 20 • The Road Ahead	 459
1. Conclusions	459
1.1 The Challenge of Assessing Impacts in a World in Which Many Evaluations Have a Short-Term Focus	459
1.2 The Continuing Debate on the “Best” Evaluation Methodologies	459
1.3 Selecting the Appropriate Evaluation Design	461
1.4 Mixed Methods: The Approach of Choice for Most RealWorld Evaluations	463
1.5 How Does RealWorld Evaluation Fit Into the Picture?	463
1.6 Quality Assurance	463
1.7 Need for a Strong Focus on Gender Equality and Social Equity	464
1.8 Basing the Evaluation Design on a Program Theory Model	465
1.9 The Importance of Context	465
1.10 The Importance of Process	466
1.11 Dealing With Complexity in Development Evaluation	466
1.12 Emergence	467
1.13 Integrating the New Information Technologies Into Evaluation	468
1.14 Greater Attention Must Be Given to the Management of Evaluations	468
1.15 The Challenge of Institutionalization of Evaluation	469
1.16 The Importance of Competent Professional and Ethical Practice	470
2. Recommendations	470
2.1 Developing Standardized Methodologies for the Evaluation of Complex Programs	470
2.2 Creative Approaches for the Definition and Use of Counterfactuals	471
2.3 Strengthening Quality Assurance and Threats to Validity Analysis	471

2.4 Defining Minimum Acceptable Quality Standards for Conducting Evaluations Under Constraints	471
2.5 Further Refinements to Program Theory	472
2.6 Further Refinements to Mixed-Method Designs	472
2.7 Integrating Big Data and Data Science Into Program Evaluation	472
2.8 Further Work Is Required to Strengthen the Integration of a Gender-Responsive Approach Into Evaluation Programs	473
2.9 The Road Ahead	473
 Glossary of Terms and Acronyms	 474
References	487
Author Index	504
Subject Index	510

LIST OF BOXES, FIGURES, AND TABLES

BOXES

Box 1.1	The Implications of New Information Technology for RealWorld Evaluation	5
Box 1.2	A Familiar Evaluation Story	9
Box 1.3	Budget and Time Constraints Have Different Implications for the Evaluation Design	11
Box 1.4	Problems in Capturing Information From or About Women	12
Box 1.5	Political Influence on the Evaluation of a Hydroelectric Program in Asia	13
Box 1.6	How Organizational and Administrative Considerations Can Affect Budget, Time, Data, and Political Constraints	14
Box 1.7	Getting Inside the “Black Box”	15
Box 2.1	Boundaries and Boundary Choices	20
Box 2.2	Defining Outcomes and Impacts	29
Box 2.3	The Main Kinds of Evaluation Design	31
Box 3.1	Six Questions to Help Make the Budget Go Farther	42
Box 3.2	“Robust” Designs Require Correct Implementation!	44
Box 3.3	Examples of Economical Methods of Data Collection	52
Box 4.1	Two Kinds of Time Saving	59
Box 4.2	Time Constraints Affecting the Evaluation of a Water and Sanitation Project in Mozambique	59
Box 4.3	Examples of Rapid Data-Collection Methods	65
Box 4.4	Rapid Ethnographic Methods (RAP)	66
Box 4.5	Reducing the Time Required to Conduct a Mixed-Method Study of Survival Strategies of Low-Income Households	66
Box 5.1	Examples of Sampling Frames That May Exclude Sectors of the Intended Target Population	75
Box 5.2	Using Administrative Data From the Project to Reconstruct the Baseline	78
Box 5.3	Examples From Developing Countries of Databases That Do Not Adequately Represent Both Sexes	79
Box 5.4	Ghost Schools and Missing Girls: Why the Validity of Secondary Data Must Always Be Checked	80
Box 5.5	Examples of Retrospective Surveys Used to Reconstruct Baseline Data	82
Box 5.6	Selecting a Wide Range of Informants When Studying Sensitive Topics: Studying Crack Cocaine Addicts in New York	83
Box 5.7	Constructing a PRA Seasonal Calendar to Reconstruct Periods of Stress in a Rural Community in Kenya	85
Box 5.8	Potentially Strong and Weak Comparison Groups	87
Box 5.9	Examples of Collecting Data on Sensitive Topics	89

Box 6.1	Example of Difficulty in Obtaining Stakeholder Input	97
Box 6.2	Observing Staged Events and Exemplary Sites	101
Box 6.3	Spinning the Findings: An Old Art Form	102
Box 7.1	Factors Affecting the Validity of CARE International's Program Evaluation Methodology	107
Box 7.2	Ways to Address Typical Threats to Statistical, Internal Design, and Construct and External Validity for a Mixed-Method Evaluation	117
Box 7.3	Frick's Process-Oriented Approach to Assessing the Quality of Qualitative Evaluations	120
Box 8.1	How Are Evaluations Used? When Are They Influential?	129
Box 8.2	Political and Bureaucratic Challenges Affecting Use of Evaluations	132
Box 8.3	Nine Actions to Promote Evaluation Utilization	133
Box 8.4	Guidelines for Helping Clients Use the Evaluation	134
Box 8.5	Enhancing Evaluation Utilization—Lessons Learned	135
Box 8.6	Using Feedback From Process Analysis to Improve Project Performance (Formative Evaluation or Utilization-Focused Evaluation)	137
Box 8.7	An Eight-Step Strategy for Communicating the Findings of Social Science Research to Policymakers	140
Box 10.1	Two Approaches to the Theory of Change	157
Box 10.2	Anticipated Benefits of Working With Theory of Change	160
Box 10.3	Behavioral Economics as a Tool for Modeling Mechanisms of Change in TOC	166
Box 10.4	Six Common Challenges Facing the Articulation of Theories of Change	171
Box 10.5	Resources for Designing a Theory of Change	174
Box 10.6	Nine Potential Types of Evidence for Inferring Causality	182
Box 11.1	The Types of Evaluation Design That Will Be Covered in This and Subsequent Chapters	189
Box 11.2	Using Contribution Analysis to Assess the Effects of UNESCO Prizes	192
Box 11.3	Examples of Methodologically Strong Nonexperimental Designs	204
Box 11.4	How Program Theory Can Strengthen an Evaluation Design	208
Box 11.5	Examples of Big Data Tools and Techniques That Might Be Available to Strengthen the RWE Design	212
Box 11.6	Building Evidence in Challenging Contexts	214
Box 12.1	The Benefits of Randomized Designs	222
Box 12.2	Methods for Randomly Assigning Subjects to the Project and Control Groups	222
Box 12.3	Challenges in Using Randomized Control Trials	223
Box 13.1	Participant Observation in a Peer Mediation Program	249
Box 13.2	Group Interviews About Women's Use of Public Transportation	252
Box 13.3	Examples of Requests in Comprehensive Validation	255
Box 14.1	Elements of an Integrated, Multidisciplinary Research Approach	282
Box 15.1	Small Samples May Overlook Important Real Effects	291
Box 15.2	Be Cautious About Advice From Statisticians Not Familiar With the Subject Area	298

Box 15.3	Longitudinal Panel Studies Are Great—But What If You Can’t Locate the Original Sampled Households Two Years Later?	302
Box 15.4	Examples of How Big Data Can Strengthen Sample Designs	304
Box 16.1	A Hypothetical Case Study to Illustrate the Different Stages of the Evaluation Framework	332
Box 16.2	Using Satellite and Remote Sensor Images to Strengthen Sample Design for Evaluating the Impacts on Forest Cover of a Forest Protection Program	343
Box 16.3	Unpacking and Reassembling Using Qualitative Comparative Analysis	346
Box 17.1	Estimated Economic Costs of Gender Inequality: Some Examples From Africa, the Middle East, and Asia	354
Box 17.2	Oxfam’s Seven Drivers of Women’s Economic Empowerment	360
Box 17.3	Elements of a Gender Analysis Framework	362
Box 17.4	Six Fundamental Beliefs of Feminist Approaches	363
Box 17.5	A Gender Adaptation of the OECD/DAC Evaluation Criteria	367
Box 17.6	Institutional Intervention Points	368
Box 17.7	Case Studies of Gender Outcomes	369
Box 17.8	Examples of Gender Checklists	371
Box 17.9	Gender Equality Integration Into SDGs	375
Box 17.10	Methods for Gender Analysis Evaluation	377
Box 17.11	Methods for Feminist Evaluation	377
Box 18.1	We Are All Leaving Traceable Digital Trails of Our Lives	384
Box 18.2	Examples of Big Data Sources That Are, or Could Be, Used for Evaluation	386
Box 18.3	Top Finalists of the World Bank Big Data Innovation Challenge	397
Box 18.4	Using an Integrated Data Platform to Help End Modern Slavery	398
Box 18.5	Potential Use of Predictive Analytics Using Data From Government Flagship Programs in Developing Countries	399
Box 18.6	Applications of Machine Learning to Evaluation	400
Box 18.7	An Example of Social Media Analytics: Rapid Assessment of Communicable Disease Control Programs in Indonesia Based on the Analysis of Tweets	404
Box 18.8	Examples of the Application of Data Science Tools in Evaluation	417
Box 18.9	Using Drones in the Evaluation of an Irrigation Project	418
Box 20.1	Programs Can Dramatically Change During Implementation: The Example of a School Feeding Program	467

FIGURES

Figure 1.1	The RealWorld Evaluation Approach	7
Figure 2.1	Step 1: Scoping the Evaluation	21
Figure 2.2	A Simple Program Theory Model	27
Figure 3.1	Step 2: Addressing Budget and Other Resource Constraints	40
Figure 4.1	Step 3: Addressing Time Constraints	56
Figure 5.1	Step 4: Addressing Data Constraints	73

Figure 5.2	When Should the Baseline Study Be Conducted for the Evaluation of a Slum-Upgrading Project?	77
Figure 6.1	Step 5: Addressing Political Influences	96
Figure 7.1	Step 6: Strengthening the Evaluation Design and the Validity of Conclusions	106
Figure 7.2	Dimensions of Threats to Validity Used to Assess Quantitative and Qualitative Evaluation Designs	110
Figure 7.3	Dimensions of Threats to Validity for the Assessment of Mixed-Method Evaluation Designs	113
Figure 7.4	Structure, Audience, and Responsibility for the Different Parts of the Threats-to-Validity Worksheet	124
Figure 8.1	Step 7: Helping Clients Use the Evaluation	128
Figure 10.1	Theory of Change as a Component of a Program Theory	155
Figure 10.2	The Theory of Change as the “Missing Middle” of the Logic Model	159
Figure 10.3	Weaknesses of the “Black Box” Approach to Evaluation	159
Figure 10.4	Using a Results Chain to Identify Potential Benefits and Unintended Negative Outcomes of a Women’s Microcredit Program	161
Figure 10.5	The Elements of the Realist Evaluation Approach	164
Figure 10.6	A Theory of Change for Tackling Violence Against Women and Girls	173
Figure 10.7	Example of Process Tracing: Hypothetical Example Tracing the Effectiveness of Mechanisms to Strengthen the Quality of Evaluation Design on the Quality, and Accessibility to Vulnerable Groups, of Social Policies and Programs	178
Figure 10.8	Different Scenarios for the Variation of Project Effects Over Time	183
Figure 11.1	Eight Steps for the Selection of the Appropriate Evaluation Design to Address Different Evaluation Contexts, Purposes, and Scenarios	188
Figure 13.1	<i>Observation Protocol: Are Ethnic Minorities Excluded From Program Benefits?</i>	247
Figure 13.2	<i>Observation Protocol: Are Low Ability-Tracked Students Inappropriately Prevented From Moving to High Tracks?</i>	247
Figure 13.3	<i>Observation Protocol: Does Provision of Laptop Computers to Students Improve Achievement, Personalization, or Rigor of Their Educations or Classroom Community?</i>	248
Figure 13.4	Guidelines for Developing Questions for Interview Protocols	252
Figure 13.5	Documents and Artifacts That May Be Available From Evaluation Sites	254
Figure 14.1	Using a National Household Sample Frame to Ensure the Representativity and Credibility of a QUANT Case Study Evaluation Design Assessing the Access of Women to Rural Health Centers	265
Figure 14.2	Using QUAL Ethnographic Village Studies to Help Interpret the Findings of a National QUANT Sample Survey Estimating Women’s Utilization Rates of Rural Health Centers	266
Figure 14.3	The QUANT–QUAL Research Continuum	270
Figure 14.4	Sequential Mixed-Method Design With a Dominant Quantitative Approach: Studying Interhousehold Transfers of Money and Goods in Cartagena, Colombia	277
Figure 14.5	Sequential Mixed-Method Design With a Dominant Qualitative Approach: Evaluating the Adoption of New Seed Varieties by Different Types of Rural Families	278

Figure 14.6	Concurrent Triangulation Design: Comparing and Integrating Different Data Sources on Poverty and Household Income	279
Figure 14.7	Collaborative Outcome Reporting Technique (CORT): Data Collection, Analysis, and Report Structure	285
Figure 15.1	A Multistage Sample for Selecting a Sample of Secondary School Students	300
Figure 16.1	The Dimensions of Complexity of Development Programs and Their Evaluation	318
Figure 16.2	Simple, Linear Causal Chain	320
Figure 16.3	Complex Causal Chain	321
Figure 16.4	Examples of Complex, Nonproportional Change	322
Figure 16.5	Simple Projects, Complicated Programs, and Complex Interventions	323
Figure 16.6	The Five-Step Unpacking-Reassembling Approach for Evaluating Complex Programs	334
Figure 16.7	Unpacking the Micro-Grant Component of a Multidonor Decentralized Transport Support Program	335
Figure 16.8	Systems Map for the Grants Component of the Ministry of Transport Project	348
Figure 16.9	Disaggregating National Programs Into National, Regional, and Local Levels	349
Figure 17.1	The Ladder of Gendered Evaluation	364
Figure 17.2	A Continuum of Gendered Evaluation	375
Figure 18.1	Elements of the New Information Technology System	385
Figure 18.2	The Data Continuum	392
Figure 18.3	The Big Data Ecosystem and the Linkages to the Evaluation Ecosystem	395
Figure 18.4	The Four Stages of the Data Analytics Cycle	400
Figure 19.1	Managing the Evaluation Process	432
Figure 19.2	Three Pathways for the Evolution of Institutionalized IE Systems	452

TABLES

Table 1.1	RealWorld Evaluation Scenarios: Conducting Impact Evaluations With Time, Budget, Data, and Political and Organizational Constraints	10
Table 1.2	Who Uses RWE, for What Purposes, and When?	17
Table 2.1	Evaluation Design Components to Strengthen All of the Basic Evaluation Designs	33
Table 2.2	Determining Possible Evaluation Designs	34
Table 3.1	Reducing Costs of Data Collection and Analysis for Quantitative and Qualitative Evaluations	41
Table 3.2	Overview of Most Common RWE Impact Evaluation Designs Classified by Their Statistical Strength	45
Table 3.3	Estimated Cost Savings for Less Robust RWE Designs Compared With Design 2.2	46

Table 3.4	Sample Size Required for Various Minimum Detectable Effects (Decrease in Household Health Expenditures), Power = 0.9. No Clustering	49
Table 4.1	Options for Reducing the Time Required for Data Collection and Analysis in Quantitative and Qualitative Evaluations	57
Table 4.2	Rapid Data-Collection Methods: The Effect on Duration of the Evaluation and the Level of Effort	63
Table 5.1	Summary of Strategies for Addressing Data Constraints	91
Table 6.1	Points at Which Political Influences May Affect an Evaluation	95
Table 7.1	Factors Determining the Adequacy of the Evaluation Design and of the Findings	108
Table 10.1	Examples of the Application of TBE at Different Stages of the Evaluation Cycle	163
Table 10.2	The Different Ways in Which a Theory of Change Can Be Used	168
Table 10.3	Variations in the Estimated Level of Effects Depending on the Year in Which the Posttest Study Is Conducted	182
Table 11.1	The Contribution Analysis Six-Step Approach	191
Table 11.2	Evaluation Purpose and Context	194
Table 11.3	Seven Basic Impact Evaluation Designs and Two More Advanced Options	199
Table 11.4	Tools and Techniques to Strengthen Any of the Basic Impact Evaluation Designs	206
Table 12.1	Strengths and Weaknesses of QUANT Evaluation Approaches	227
Table 12.2	Features of Well-Designed Quantitative Evaluations	236
Table 14.1	Strengths and Weaknesses of Qualitative Evaluation Approaches	264
Table 14.2	Examples of Evaluation Designs at Each Point on the QUANT–QUAL Continuum	271
Table 15.1	Approximate Required Sample Size for Comparing a Single Sample Mean With the Population Mean to Achieve Different Levels of Power for a Range of Effect Sizes (With a One-Tail Test and Significance Level = 0.05)	308
Table 15.2	Mean Effect Sizes for Different Interventions Addressing Aggressive Behavior in Schools	309
Table 15.3	Estimated Minimum Sample Sizes Required for Different Types of Qualitative Research Components of Mixed-Method Designs	311
Table 15.4	Lot Quality Acceptance Sampling: Decision Rules for Determining the Minimum Number of Units With Satisfactory Delivery of Information or Services to Accept That a Particular Benchmark Level of Satisfactory Services Has Been Achieved (β errors > 10%)	313
Table 16.1	Complicated and Complex Aspects of Interventions	326
Table 16.2	Checklist for Assessing the Level of Complexity of a Program	328
Table 16.3	The Potential Applications of Complexity Science in the Complexity-Responsive Evaluation Approach (Framework)	336
Table 16.4	Alternative Approaches for Defining the Counterfactual With Complex Program Evaluations	341

Table 17.1	Scenarios for the Integration of Gender Into Evaluation	367
Table 18.1	Defining Big Data	388
Table 18.2	Comparing Big Data and Conventional Evaluation Data	389
Table 18.3	Widely Used NIT and Data Analytics Applications in International Development	396
Table 18.4	Kinds of Big Data Analysis With Potential Applications for Program Monitoring and Evaluation	402
Table 18.5	Some Widely Cited Categories of Data Analytics (Data Science)	407
Table 18.6	Potential Ways That Big Data and ICTs Can Strengthen Program Evaluation	413
Table 18.7	Examples Illustrating How Big Data and ICTs Have Been Used to Strengthen Widely Used Evaluation Designs	416
Table 19.1	Roles of Stakeholders in an Evaluation	434
Table 19.2	How Lack of Coordination Between Project Management and the Evaluator Can Weaken the Evaluation Design	443