

Program Evaluation and Performance Measurement

An Introduction to Practice

Third Edition

James C. McDavid

University of Victoria, Canada

Irene Huse

University of Victoria, Canada

Laura R. L. Hawthorn



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

CONTENTS

Preface	xv
Acknowledgments	xvii
About the Authors	1
Chapter 1 • Key Concepts and Issues in Program Evaluation and Performance Management	2
Introduction	3
Integrating Program Evaluation and Performance Measurement	4
Connecting Evaluation to the Performance Management System	5
The Performance Management Cycle	8
Policies and Programs	10
Key Concepts in Program Evaluation	12
Causality in Program Evaluations	12
Formative and Summative Evaluations	14
Ex Ante and Ex Post Evaluations	15
The Importance of Professional Judgment in Evaluations	16
Example: Evaluating a Police Body-Worn Camera Program in Rialto, California	17
The Context: Growing Concerns With Police Use of Force and Community Relationship	17
Implementing and Evaluating the Effects of Body-Worn Cameras in the Rialto Police Department	18
Program Success Versus Understanding the Cause-and-Effect Linkages: The Challenge of Unpacking the Body-Worn Police Cameras “Black Box”	20
Connecting Body-Worn Camera Evaluations to This Book	21
Ten Key Evaluation Questions	22
The Steps in Conducting a Program Evaluation	28
General Steps in Conducting a Program Evaluation	28
<i>Assessing the Feasibility of the Evaluation</i>	30
<i>Doing the Evaluation</i>	37
<i>Making Changes Based on the Evaluation</i>	41
Summary	43
Discussion Questions	44
References	45
Chapter 2 • Understanding and Applying Program Logic Models	50
Introduction	51
Logic Models and the Open Systems Approach	52
A Basic Logic Modeling Approach	54
An Example of the Most Basic Type of Logic Model	58
Working With Uncertainty	60
Problems as Simple, Complicated, and Complex	60
Interventions as Simple, Complicated, or Complex	61
The Practical Challenges of Using Complexity Theory in Program Evaluations	62

Program Objectives and Program Alignment With Government Goals	64
Specifying Program Objectives	64
Alignment of Program Objectives With Government and Organizational Goals	66
Program Theories and Program Logics	68
Systematic Reviews	69
Contextual Factors	70
Realist Evaluation	71
Putting Program Theory Into Perspective: Theory-Driven Evaluations and Evaluation Practice	74
Logic Models That Categorize and Specify Intended Causal Linkages	75
Constructing A Logic Model For Program Evaluations	79
Logic Models For Performance Measurement	81
Strengths and Limitations of Logic Models	84
Logic Models in a Turbulent World	85
Summary	86
Discussion Questions	87
Appendices	88
Appendix A: Applying What You Have Learned: Development of a Logic Model for a Meals on Wheels Program	88
<i>Translating a Written Description of a Meals on Wheels Program Into a Program Logic Model</i>	88
Appendix B: A Complex Logic Model Describing Primary Health Care in Canada	88
Appendix C: Logic Model for the Canadian Evaluation Society Credentialed Evaluator Program	92
References	94
Chapter 3 • Research Designs For Program Evaluations	97
Introduction	98
Our Stance	98
What Is Research Design?	104
The Origins of Experimental Design	105
Why Pay Attention to Experimental Designs?	110
Using Experimental Designs to Evaluate Programs	112
The Perry Preschool Study	112
Limitations of the Perry Preschool Study	115
The Perry Preschool Study in Perspective	116
Defining and Working With the Four Basic Kinds of Threats to Validity	118
Statistical Conclusions Validity	118
Internal Validity	118
Police Body-Worn Cameras: Randomized Controlled Trials and Quasi-Experiments	122
Construct Validity	124
The 'Measurement Validity' Component of Construct Validity	125
<i>Other Construct Validity Problems</i>	126
External Validity	129
Quasi-experimental Designs: Navigating Threats to Internal Validity	131
The York Neighborhood Watch Program: An Example of an Interrupted Time Series	136
Research Design Where the Program Starts, Stops, and Then Starts Again	136
Findings and Conclusions From the Neighborhood Watch Evaluation	137
Non-Experimental Designs	140

Testing the Causal Linkages in Program Logic Models	141
Research Designs and Performance Measurement	145
Summary	147
Discussion Questions	148
Appendices	150
Appendix 3A: Basic Statistical Tools for Program Evaluation	150
Appendix 3B: Empirical Causal Model for the Perry Preschool Study	152
Appendix 3C: Estimating the Incremental Impact of a Policy Change—Implementing and Evaluating an Admission Fee Policy in the Royal British Columbia Museum	153
References	157
Chapter 4 • Measurement for Program Evaluation and Performance Monitoring	161
Introduction	162
Introducing Reliability and Validity of Measures	164
Understanding the Reliability of Measures	167
Understanding Measurement Validity	169
Types of Measurement Validity	170
Ways to Assess Measurement Validity	171
<i>Validity Types That Relate a Single Measure to a Corresponding Construct</i>	172
<i>Validity Types That Relate Multiple Measures to One Construct</i>	172
<i>Validity Types That Relate Multiple Measures to Multiple Constructs</i>	173
Units of Analysis and Levels of Measurement	175
Nominal Level of Measurement	176
Ordinal Level of Measurement	177
Interval and Ratio Levels of Measurement	177
Sources of Data in Program Evaluations and Performance Measurement Systems	179
Existing Sources of Data	179
Sources of Data Collected by the Program Evaluator	182
Surveys as an Evaluator-Initiated Data Source in Evaluations	182
<i>Working With Likert Statements in Surveys</i>	185
<i>Designing and Conducting Surveys</i>	187
<i>Structuring Survey Instruments: Design Considerations</i>	189
Using Surveys to Estimate the Incremental Effects of Programs	192
Addressing Challenges of Personal Recall	192
Retrospective Pre-tests: Where Measurement Intersects With Research Design	194
Survey Designs Are Not Research Designs	196
Validity of Measures and the Validity of Causes and Effects	197
Summary	199
Discussion Questions	201
References	202
Chapter 5 • Applying Qualitative Evaluation Methods	205
Introduction	206
Comparing and Contrasting Different Approaches To Qualitative Evaluation	207
Understanding Paradigms and Their Relevance to Evaluation	208
Pragmatism as a Response to the Philosophical Divisions Among Evaluators	213
Alternative Criteria for Assessing Qualitative Research and Evaluations	214

Qualitative Evaluation Designs: Some Basics	216
Appropriate Applications for Qualitative Evaluation Approaches	216
Comparing and Contrasting Qualitative and Quantitative Evaluation Approaches	218
Designing and Conducting Qualitative Program Evaluations	221
1. Clarifying the Evaluation Purpose and Questions	222
2. Identifying Research Designs and Appropriate Comparisons	222
<i>Within-Case Analysis</i>	222
<i>Between-Case Analysis</i>	223
3. Mixed-Methods Evaluation Designs	224
4. Identifying Appropriate Sampling Strategies in Qualitative Evaluations	228
5. Collecting and Coding Qualitative Data	230
<i>Structuring Data Collection Instruments</i>	230
<i>Conducting Qualitative Interviews</i>	231
6. Analyzing Qualitative Data	233
7. Reporting Qualitative Results	237
Assessing the Credibility and Generalizability of Qualitative Findings	237
Connecting Qualitative Evaluation Methods to Performance Measurement	239
The Power of Case Studies	241
Summary	243
Discussion Questions	244
References	245
Chapter 6 • Needs Assessments for Program Development and Adjustment	248
Introduction	249
General Considerations Regarding Needs Assessments	250
<i>What Are Needs and Why Do We Conduct Needs Assessments?</i>	250
<i>Group-Level Focus for Needs Assessments</i>	252
<i>How Needs Assessments Fit Into the Performance Management Cycle</i>	252
<i>Recent Trends and Developments in Needs Assessments</i>	254
<i>Perspectives on Needs</i>	255
<i>A Note on the Politics of Needs Assessment</i>	256
Steps in Conducting Needs Assessments	257
Phase I: Pre-Assessment	259
1. <i>Focusing the Needs Assessment</i>	260
2. <i>Forming the Needs Assessment Committee (NAC)</i>	266
3. <i>Learning as Much as We Can About Preliminary "What Should Be" and "What Is" Conditions From Available Sources</i>	267
4. <i>Moving to Phase II and/or III or Stopping</i>	268
Phase II: The Needs Assessment	268
5. <i>Conducting a Full Assessment About "What Should Be" and "What Is"</i>	268
6. <i>Needs Assessment Methods Where More Knowledge Is Needed: Identifying the Discrepancies</i>	269
7. <i>Prioritizing the Needs to Be Addressed</i>	278
8. <i>Causal Analysis of Needs</i>	280
9. <i>Identification of Solutions: Preparing a Document That Integrates Evidence and Recommendations</i>	280
10. <i>Moving to Phase III or Stopping</i>	282
Phase III: Post-Assessment: Implementing a Needs Assessment	283
11. <i>Making Decisions to Resolve Needs and Select Solutions</i>	283
12. <i>Developing Action Plans</i>	284
13. <i>Implementing, Monitoring and Evaluating</i>	284

Needs Assessment Example: Community Health Needs Assessment in New Brunswick	285
The Needs Assessment Process	286
<i>Focusing the Needs Assessment</i>	286
<i>Forming the Needs Assessment Committee</i>	286
<i>Learning About the Community Through a Quantitative Data Review</i>	287
<i>Learning About Key Issues in the Community Through Qualitative Interviews and Focus Groups</i>	288
<i>Triangulating the Qualitative and Quantitative Lines of Evidence</i>	288
<i>Prioritizing Primary Health-Related Issues in the Community</i>	288
Summary	291
Discussion Questions	292
Appendixes	293
Appendix A: Case Study: Designing a Needs Assessment for a Small Nonprofit Organization	293
<i>The Program</i>	293
<i>Your Role</i>	294
<i>Your Task</i>	294
References	295
Chapter 7 • Concepts and Issues in Economic Evaluation	298
Introduction	299
Why an Evaluator Needs to Know About Economic Evaluation	300
Connecting Economic Evaluation With Program Evaluation: Program Complexity and Outcome Attribution	302
<i>Program Complexity and Determining Cost-Effectiveness of Program Success</i>	302
<i>The Attribution Issue</i>	303
Three Types of Economic Evaluation	304
The Choice of Economic Evaluation Method	304
Economic Evaluation in the Performance Management Cycle	306
Historical Developments in Economic Evaluation	307
Cost-Benefit Analysis	308
Standing	309
Valuing Nonmarket Impacts	312
<i>Revealed and Stated Preferences Methods for Valuing Nonmarket Impacts</i>	312
Steps for Economic Evaluations	313
1. <i>Specify the Set of Alternatives</i>	314
2. <i>Decide Whose Benefits and Costs Count (Standing)</i>	314
3. <i>Categorize and Catalog the Costs and Benefits</i>	314
4. <i>Predict Costs and Benefits Quantitatively Over the Life of the Project</i>	315
5. <i>Monetize (Attach Dollar Values to) All Costs and Benefits</i>	315
6. <i>Select a Discount Rate for Costs and Benefits Occurring in the Future</i>	316
7. <i>Compare Costs With Outcomes, or Compute the Net Present Value of Each Alternative</i>	317
8. <i>Perform Sensitivity and Distributional Analysis</i>	318
9. <i>Make a Recommendation</i>	319
Cost-Effectiveness Analysis	320
Cost-Utility Analysis	321
Cost-Benefit Analysis Example: The High/Scope Perry Preschool Program	322
1. Specify the Set of Alternatives	324
2. Decide Whose Benefits and Costs Count (Standing)	324
3. Categorize and Catalog Costs and Benefits	324
4. Predict Costs and Benefits Quantitatively Over the Life of the Project	325

5. Monetize (Attach Dollar Values to) All Costs and Benefits	325
6. Select a Discount Rate for Costs and Benefits Occurring in the Future	326
7. Compute the Net Present Value of the Program	327
8. Perform Sensitivity and Distributional Analysis	327
9. Make a Recommendation	327
Strengths and Limitations of Economic Evaluation	328
Strengths of Economic Evaluation	328
Limitations of Economic Evaluation	329
Summary	333
Discussion Questions	334
References	336
Chapter 8 • Performance Measurement as an Approach to Evaluation	340
Introduction	341
The Current Imperative To Measure Performance	342
Performance Measurement For Accountability and Performance Improvement	343
Growth and Evolution of Performance Measurement	344
Performance Measurement Beginnings in Local Government	344
Federal Performance Budgeting Reform	345
The Emergence of New Public Management	346
Steering, Control, and Performance Improvement	349
Metaphors That Support and Sustain Performance Measurement	350
Organizations as Machines	351
Government as a Business	351
Organizations as Open Systems	352
Comparing Program Evaluation and Performance Measurement Systems	353
Summary	364
Discussion Questions	365
References	366
Chapter 9 • Design and Implementation of Performance Measurement Systems	371
Introduction	372
The Technical/Rational View and the Political/Cultural View	372
Key Steps in Designing and Implementing a Performance Measurement System	374
1. Leadership: Identify the Organizational Champions of This Change	375
2. Understand What Performance Measurement Systems Can and Cannot Do	377
3. Communication: Establish Multi-Channel Ways of Communicating That Facilitate Top-Down, Bottom-Up, and Horizontal Sharing of Information, Problem Identification, and Problem Solving	379
4. Clarify the Expectations for the Intended Uses of the Performance Information That is Created	380
5. Identify the Resources and Plan for the Design, Implementation, and Maintenance of the Performance Measurement System	383
6. Take the Time to Understand the Organizational History Around Similar Initiatives	384
7. Develop Logic Models for the Programs for Which Performance Measures Are Being Designed and Identify the Key Constructs to Be Measured	385
8. Identify Constructs Beyond Those in Single Programs: Consider Programs Within Their Place in the Organizational Structure	387

9. Involve Prospective Users in Development of Logic Models and Constructs in the Proposed Performance Measurement System	390
10. Translate the Constructs Into Observable Performance Measures that Compose the Performance Measurement System	391
11. Highlight the Comparisons That Can Be Part of the Performance Measurement System	395
12. Reporting and Making Changes to the Performance Measurement System	398
Performance Measurement for Public Accountability	400
Summary	402
Discussion Questions	403
Appendix A: Organizational Logic Models	404
References	405
Chapter 10 • Using Performance Measurement for Accountability and Performance Improvement	409
Introduction	410
Using Performance Measures	411
Performance Measurement in a High-Stakes Environment: The British Experience	412
<i>Assessing the “Naming and Shaming” Approach to Performance Management in Britain</i>	415
<i>A Case Study of Gaming: Distorting the Output of a Coal Mine</i>	418
Performance Measurement in a Medium-Stakes Environment: Legislator Expected Versus Actual Uses of Performance Reports in British Columbia, Canada	419
The Role of Incentives and Organizational Politics in Performance Measurement Systems With a Public Reporting Emphasis	424
Performance Measurement in a Low-Stakes Environment: Joining Internal and External Uses of Performance Information in Lethbridge, Alberta	425
Rebalancing Accountability-Focused Performance Measurement Systems to Increase Performance Improvement Uses	429
Making Changes to a Performance Measurement System	432
Does Performance Measurement Give Managers the “Freedom to Manage?”	434
Decentralized Performance Measurement: The Case of a Finnish Local Government	435
When Performance Measurement Systems De-Emphasize Outputs and Outcomes: Performance Management Under Conditions of Chronic Fiscal Restraint	437
Summary	439
Discussion Questions	440
References	441
Chapter 11 • Program Evaluation and Program Management	445
Introduction	446
Internal Evaluation: Views From the Field	447
Intended Evaluation Purposes and Managerial Involvement	450
<i>When the Evaluations Are for Formative Purposes</i>	450
<i>When the Evaluations Are for Summative Purposes</i>	452
Optimizing Internal Evaluation: Leadership and Independence	453
<i>Who Leads the Internal Evaluation?</i>	454
<i>“Independence” for Evaluators</i>	455
Building an Evaluative Culture in Organizations: An Expanded Role for Evaluators	456
Creating Ongoing Streams of Evaluative Knowledge	457

Critical Challenges to Building and Sustaining an Evaluative Culture	458
Building an Evaluative/Learning Culture in a Finnish Local Government:	
Joining Performance Measurement and Performance Management	460
Striving for Objectivity in Program Evaluations	460
Can Program Evaluators Claim Objectivity?	462
Objectivity and Replicability	463
<i>Implications for Evaluation Practice: A Police Body-Worn Cameras Example</i>	466
Criteria for High-Quality Evaluations	467
Summary	470
Discussion Questions	471
References	472
Chapter 12 • The Nature and Practice of Professional Judgment in Evaluation	477
Introduction	478
The Nature of the Evaluation Enterprise	478
Our Stance	479
Reconciling the Diversity in Evaluation Theory With Evaluation Practice	480
Working in the Swamp: The Real World of Evaluation Practice	481
Ethical Foundations of Evaluation Practice	482
Power Relationships and Ethical Practice	485
Ethical Guidelines for Evaluation Practice	486
Evaluation Association-Based Ethical Guidelines	486
Understanding Professional Judgment	490
What Is Good Evaluation Theory and Practice?	490
Tacit Knowledge	492
Balancing Theoretical and Practical Knowledge in Professional Practice	492
Aspects of Professional Judgment	493
The Professional Judgment Process: A Model	495
<i>The Decision Environment</i>	497
<i>Values, Beliefs, and Expectations</i>	497
Cultural Competence in Evaluation Practice	498
Improving Professional Judgment in Evaluation	499
Mindfulness and Reflective Practice	499
Professional Judgment and Evaluation Competencies	501
Education and Training-Related Activities	504
Teamwork and Improving Professional Judgment	505
The Prospects for an Evaluation Profession	506
Summary	509
Discussion Questions	510
Appendix	511
Appendix A: Fiona’s Choice: An Ethical Dilemma for a Program Evaluator	511
Your Task	512
References	513
Glossary	517
Index	530