8-2008

Evaluating Organizational Effectiveness

Wesley A. Martz
Western Michigan University

Follow this and additional works at: https://scholarworks.wmich.edu/dissertations

Part of the Policy Design, Analysis, and Evaluation Commons, and the Public Administration Commons

Recommended Citation
https://scholarworks.wmich.edu/dissertations/793

This Dissertation-Open Access is brought to you for free and open access by the Graduate College at ScholarWorks at WMU. It has been accepted for inclusion in Dissertations by an authorized administrator of ScholarWorks at WMU. For more information, please contact maira.bundza@wmich.edu.
EVALUATING ORGANIZATIONAL EFFECTIVENESS

by

Wesley A. Martz

A Dissertation
Submitted to the
Faculty of The Graduate College
in partial fulfillment of the
requirements for the
Degree of Doctor of Philosophy
Department of Interdisciplinary Ph.D. in Evaluation
Dr. K.C. O'Shaughnessy, Advisor

Western Michigan University
Kalamazoo, Michigan
August 2008
ACKNOWLEDGMENTS

My first inquiry into the Interdisciplinary Ph.D. Program in Evaluation at Western Michigan University brought me into contact with Dr. Jane Davidson, then director of the program. Spending the previous fifteen years in private industry, I was not quite sure the program was applicable for a business professional seeking to continue work in business and industry. However, Dr. Davidson’s enthusiastic and affirmative reply to my questions encouraged me to apply to the program. My acceptance letter welcoming me to the program was received from the “new” director of the program, who also served as my advisor throughout my doctoral studies, Dr. Michael Scriven. It is with great respect and appreciation that I thank both of these inspiring individuals for leading me into the field of evaluation.

Throughout the coursework, evaluation practicums, and special projects, I was encouraged and supported by my cohort, who became my friends: Amy Gullickson, Otto Gustafson, and Ryoh Sasaki. Thank you for your willingness to engage. Sincere thanks go to another central figure in my doctoral studies, Mary Ramlow, whose intimate knowledge of the University “made things happen.” A special mention is given to Dr. Chris Coryn for his scholarly advice, warm critique of my work, and his kindness that has led to a new friendship. I also offer sincere appreciation to Dr. Daniela Schröter who not only joined with me to form the Dissertation Support Group, but provided reality checks for me during the dissertation process. And to my special friend,
Acknowledgments—Continued

Dr. Lois-ellin Datta, your uplifting words and practical advice were an encouragement to me. Mahalo.

Three exceptional individuals deserve special mention for their willingness to serve on my dissertation committee: Drs. K.C. O'Shaughnessy, Liliana Rodriguez-Campos, and Michael Scriven. Dr. O'Shaughnessy's candid advice and desire to see me complete this journey provided motivation throughout the demanding dissertation writing process. Dr. Rodriguez-Campos provided guidance and support from the first day I entered the program, and continued to challenge, inspire, advise, and occasionally console me throughout my doctoral studies. Dr. Scriven, a caring scholar whose reputation has blazed a righteous path for evaluators around the globe, provided inspiration, guidance, motivation, and stimulation each and every time we interfaced. Thank you for believing in me.

And finally, to my parents Aaron and Judy, the words "thank you" hardly seem appropriate to convey my gratitude for your eternal encouragement, unconditional love, and endless support provided not only during my doctoral studies, but throughout my personal and professional endeavors. With adoration and sincere affection, thank you.

Wesley A. Martz
# TABLE OF CONTENTS

ACKNOWLEDGMENTS .......................................................................................................................... ii

LIST OF TABLES ................................................................................................................................. ix

LIST OF FIGURES ............................................................................................................................... x

CHAPTER

I. INTRODUCTION .............................................................................................................................. 1

   The Essence of Evaluation .............................................................................................................. 2

   Evaluation in Organizations .......................................................................................................... 4

      Level 1 Evaluation ....................................................................................................................... 5

      Level 2 Evaluation ..................................................................................................................... 6

      Level 3 Evaluation ..................................................................................................................... 7

   Value of Organizational Evaluation .......................................................................................... 9

   Barriers to Advancing Evaluation in Organizations ................................................................. 10

   Value Claims in Evaluation ......................................................................................................... 11

   Evaluation Anxiety ....................................................................................................................... 13

   Performance Measurement Proliferation ...................................................................................... 14

   Conceptualizing Organizational Effectiveness .............................................................................. 15

      Efficiency as Effectiveness ....................................................................................................... 16

      Organization Typology Perspective .......................................................................................... 17
# Table of Contents—Continued

## CHAPTER

The Complexity Issue .................................................. 18

Toward Evaluating Organizational Effectiveness ....................... 19

### II. ELEMENTS AND ISSUES IN ORGANIZATIONAL ASSESSMENT ...

Introduction to Organizations ........................................... 23

The Organic Perspective .................................................. 24

The Mechanical Perspective ............................................... 25

Aspects of Organizational Effectiveness ................................ 26

Level of Analysis .......................................................... 27

Nature of Organization .................................................... 28

Values and Organizational Effectiveness ................................ 30

Traditional Approaches to Assessing Organizational Effectiveness .... 32

Goal Model ................................................................. 33

Systems Model ............................................................. 37

Process Model ............................................................. 41

Strategic Constituencies Model .......................................... 45

Competing Values Framework ........................................... 49

Baldrige National Quality Program ..................................... 53

Summary ........................................................................ 57
Table of Contents—Continued

CHAPTER

III. A CHECKLIST APPROACH TO EVALUATING ORGANIZATIONAL EFFECTIVENESS ................................................................. 59

The Role of Evaluation Checklists ......................................................................................................................... 60

Organizational Effectiveness Evaluation Checklist ................................................................................................. 62

Outline of the OEC ..................................................................................................................................................... 66

Process Overview ....................................................................................................................................................... 68

Step 1: Establish the Boundaries of the Evaluation ................................................................................................. 68

Step 2: Conduct a Performance Needs Assessment ................................................................................................. 75

Step 3: Define the Criteria for the Evaluation .......................................................................................................... 79

Step 4: Plan and Implement the Evaluation ........................................................................................................ 91

Step 5: Synthesize Performance Data with Values ................................................................................................. 93

Step 6: Communicate and Report Evaluation Activities ....................................................................................... 97

Summary ................................................................................................................................................................. 102

IV. METHODOLOGY ................................................................................................................................................ 103

Overview ............................................................................................................................................................... 103

Expert Panel Review .............................................................................................................................................. 103

Participants .......................................................................................................................................................... 103

Instrumentation .................................................................................................................................................... 105

Procedures .......................................................................................................................................................... 107

Field Study Using the OEC .................................................................................................................................. 109
# Table of Contents—Continued

## CHAPTER

Overview ........................................................................................................ 109
Evaluation Client .............................................................................................. 110
Procedures ......................................................................................................... 111
Aspects of Reliability and Validity ....................................................................... 113

### V. EXPERT PANEL AND CASE STUDY FINDINGS .................................................. 116

Expert Panel Critical Review ................................................................................... 116
Qualitative Feedback: Narrative Critique ............................................................... 121
Observations from Survey Responses .................................................................. 126
Qualitative Feedback: Written Comments ............................................................ 128
Discussion of Expert Panel Findings ...................................................................... 130
Revisions to the Original OEC .............................................................................. 132
Results of Field Study Using the OEC ................................................................. 136
Evaluator Perspectives .......................................................................................... 136
Client Perspectives ............................................................................................... 139

### VI. CONCLUSIONS ............................................................................................ 143

Implications of the OEC ....................................................................................... 145
Professional Evaluation Practice ........................................................................... 145
Organizational Performance ................................................................................ 146
Theory and Research ............................................................................................ 147
Limitations of this Research ................................................................................ 149
Table of Contents—Continued

CHAPTER

Future Areas for Exploration ................................................................. 150

REFERENCES ..................................................................................... 152

APPENDICES ..................................................................................... 164

A. Organizational Effectiveness Evaluation Checklist (OEC) .................. 164

B. Critical Feedback Survey Instrument .............................................. 198

C. Sociodemographic Questionnaire .................................................. 201

D. Consent Document ........................................................................ 203

E. Approval Letter From the Human Subjects Institutional Review Board .................................................................................. 206

F. Permission to Use Copyright Information ....................................... 208
LIST OF TABLES

1. Baldrige National Quality Program Point Values for 2008 Criteria and Subcriteria ........................................ 55
2. Performance Matrix for Determining Absolute Merit of the Management Support Criterion ........................................ 90
3. Performance Matrix for Determining Relative Merit of the Profit Per Employee Criterion ........................................ 90
4. Critical Feedback Survey Ratings by Respondent Role ........................................ 117
5. Rank Order Criteria by Group ........................................................................... 118
6. ANOVA Summary Table for Mean Total Score Among Evaluators, Managers, and Theorists ........................................ 119
7. F-value, Significance, and Effect Size for Applicability, Concreteness, Ease of Use, Fairness, Parsimony, and Pertinence ........................................ 120
8. Welch Test of Equality of Means for Clarity and Comprehensiveness ........................................ 120
9. Representative Sample of Recording Units and Classifications ........................................ 122
10. Net Strength Score ........................................................................................................... 127
11. Sample Verbatim Written Comments and Classifications ........................................ 129
LIST OF FIGURES

1. Common Organizational Functions Found in an Open-System Environment ................................................................. 20
2. The Four Quadrants of the Competing Values Framework ................................................................. 51
3. Profile of Sustainable Dimension ........................................................................................................ 95
4. Organizational Effectiveness Profile ...................................................................................................... 96
5. Expert Panel Members' Extent of Professional Experience in Research, Evaluation, and Management ................................................................................................................................ 105
6. Number of Strengths Identified by Category and Primary Role .............................................................. 124
7. Number of Weaknesses Identified by Category and Primary Role ............................................................ 124
8. Number of Items Missing Identified by Category and Primary Role ........................................................ 125
9. Number of Improvements Suggested by Category and Primary Role ..................................................... 126
10. Net Strength and Mean Score Plot for Applicability, Clarity, Comprehensiveness, Concreteness, Ease of Use, Fairness, Parsimony, and Pertinence to the Content Area ................................................................................................................................. 128
11. Frequency of Written Comment Classifications .................................................................................... 130
CHAPTER I

INTRODUCTION

Societies in their very nature represent organized groupings of people whose activities are coordinated through institutional channels. In industrialized societies, formal organizations are ubiquitous and an indispensable artifact of modern human life. Organizational researchers have conceptualized organizations using a variety of expressions, including rational entities in pursuit of goals (Latham & Locke, 1991; Perrow, 1970); coalitions reacting to strategic constituencies (Pfeffer & Salancik, 1978); individual need-meeting cooperatives (Cummings, 1977); garbage cans (Cohen, March, & Olsen, 1972); open systems (Katz & Kahn, 1978); social contracts (Keeley, 1980); political arenas (Gaertner & Ramnarayan, 1983); psychic prisons (Morgan, 1986); and so on. Each of these conceptualizations highlights or reveals organizational phenomena that were overlooked or disregarded by other definitions. As the definition of the organization has changed, so has the definition of what an effective organization is. This fragmentation of organizational conceptualizations has left the organizational effectiveness construct in disarray and prevented cumulative efforts to develop a coherent model or framework for evaluating organizational effectiveness.

Conceptually, organizational effectiveness is an enigma. On one hand, it is one of the most central constructs studied by organizational researchers. It is the ultimate dependent variable for organization researchers and of high relevance to practitioners and those engaged with the organization. On the other hand, its definition and
conceptualization are ambiguous, and numerous perspectives have been taken to describe organizational effectiveness. Notwithstanding the meaningful attempts to clarify the concept of organizational effectiveness and advance organizational effectiveness evaluation, limited progress has been made toward reaching consensus on a single model or approach. Instead, most developments have focused on a particular conceptualization of organizational effectiveness informed by a particular characterization of the organization and the values of the relevant constituency (e.g., management, employees, consumers, owners, taxpayers, investors, directors, community members). Moreover, the various approaches ignore or downplay issues related to boundary specifications (including the domain of activity), criterion instability, weighting criteria of merit, suboptimization, and the synthesis of multiple evaluative findings. Perhaps the most concerning aspect of the established approaches is the tendency to stop short of making explicitly evaluative conclusions and leaving it to the audience to draw its own conclusions regarding the merit, worth, or significance of the evaluand.

The Essence of Evaluation

Evaluation is the systematic determination of the merit, worth, or significance of something (Scriven, 1991). That something, referred to as the evaluand, may be further classified as a program, product, person, policy, performance, proposal, or portfolio. In organization-specific terms, the evaluand may also be referred to as an activity, initiative, strategy, department, division, business unit, or the organization itself. As is the case for other transdisciplines—such as design, logic, measurement, and statistics—evaluation serves many practical purposes, and its general usage pervades society. However, its extraordinary multidisciplinarity (i.e., the use of methods from several disciplines, for
example, economics, psychology, sociology, management, and ethics) suggests that evaluation is of fundamental importance to human endeavors and may be more appropriately termed a supra-transdiscipline.

Sound evaluation allows individuals and organizations to distinguish between the worthwhile and the worthless, the good from the bad. Evaluation leads to the discovery of merit or quality (or lack thereof); supporting decisions such as which school to attend, which car to purchase, which program to develop, which charitable organization to continue funding, to which change initiatives to devote resources and which to abandon, and which strategy is the most cost-effective to pursue. It also provides insights and accountability at different points in the evaluand's lifecycles—from needs assessment through planning, quality control, risk management, and process improvement. Evaluation may also be done retrospectively (i.e., an ascriptive evaluation), generally for documentation or for historical truth seeking, rather than to support any decision. Supplementing the formative and summative decision-support orientations, ascriptive evaluation demonstrates the research role of evaluation (Scriven, 1993).

Although the intellectual roots and systematic approaches of evaluation predate social science by millennia (Scriven, 1991), the methodological roots of modern evaluation are in social research methods applied to improve a particular program (Fitzpatrick, Sanders, & Worthen, 2004). During the explosive development of evaluation in the 1960s, the programs under investigation were typically social programs in education, welfare, public works, urban development, health, and other human service areas. In the private sector, planning, control, and accountability during the same period were heavily influenced by the scientific management approach (Taylor, 1923) that focused on efficiency and systematization in the organizational setting. Internal controls,
standardized operating procedures, and compliance audits became integral components of organizational structure and processes. Accountability within organizations was driven internally by the organization’s focus on efficiency, productivity, and maximization of returns. Governance boards, shareholders, competition, and consumer choice enforced accountability of the organization externally. Accordingly, private sector organizations were held accountable to constituencies both internal and external to the organization, and survival of the organization was dependent on satisfying these constituencies. A similar type of accountability for the government-sponsored social programs was absent, and significant increases in government spending on antipoverty and other social programs following World War II prompted concerns in Congress, the media, and the public over the lack of accountability. Consequently, a congressional evaluation mandate was enacted in 1965, signaling the birth of contemporary evaluation (Fitzpatrick et al., 2004).

Evaluation in Organizations

Evaluation of one form or another is a regular activity in organizations, yet few managers or business professionals refer to their work as evaluation. Terms such as benchmarking, assessing, auditing, researching, and reviewing are used fluently within organizational settings, while evaluation is reserved primarily for referring to performance appraisals. Beyond the natural inclination for most humans—and thereby organizations—to do good work rather than bad, three basic levels of evaluation can be used to distinguish the types of evaluations performed in organizations (Scriven, 2008).
Level 1 Evaluation

The first level includes the conscious use of skilled evaluation, typically in the form of quality assurance methods or performance measurement systems. The particular methods used may range from standard operating procedures and company policies to advanced statistical modeling and performance dashboards. This basic level of evaluation also includes collecting feedback from the organization's clients regarding quality or organizational performance and the use of explicit techniques designed to scan the external environment for emerging trends that could affect the organization. Furthermore, applications of the first level of evaluation can be found in an organization's attempt at self-assessment using external referents or accreditation guidelines issued by various assessment agencies or international or national standards organizations (e.g., International Organization for Standards [ISO], International Electrotechnical Commission [IEC], American National Standards Institute [ANSI], German Institute for Standardization [DIN], Commission on Accreditation of Rehabilitation Facilities [CARF], Association to Advance Collegiate Schools of Business [AACSB]). Most contemporary organizations utilize some form of this first level of skilled evaluation, even if done on an ad-hoc basis.

This basic level of conscious evaluation use in organizations is not surprising considering the strong emphasis that organizations, including those in the public and nonprofit sectors, place on performance measurement (Ekerson, 2006; McDavid & Hawthorn, 2006). Although performance measurement and evaluation share a number of common elements (e.g., identification of metrics and measures, data collection, data analysis and interpretation), performance measurement is more appropriately categorized as a data collection and monitoring component of the evaluation process rather than an
alternative to serious evaluation. As outlined earlier, organizational effectiveness reflects a construct perspective in which the conceptualization of the organization is of major focus. In contrast, performance measurement endorses a process perspective with a focus on quantifying internal processes using a set of metrics (Henri, 2004). A number of performance measurement frameworks are used by organizations—from cybernetic models that rely on financial information for planning and control cycles to the more recent holistic approaches that take into account nonfinancial measures and are used as signaling and learning devices (Simons, 2000). The major emphasis on quality assurance and the widespread use of performance measurement tools by organizations has broadened the application of first-level evaluation in organizations and increased the potential for advancing evaluation use in the organizational environment.

Level 2 Evaluation

At the intermediate level, the organization attempts to supplement its internal evaluation systems with regular external evaluations. Here the external evaluator’s primary role is that of an auditor who provides an independent, third-party assessment to confirm or validate adherence to specific procedures and policies. For example, a financial audit consists of an examination of the organization’s financial statements by an external auditor or audit team, resulting in the publication of an independent opinion on whether or not the statements are relevant, accurate, and complete. No conclusions are made regarding the merit of financial performance or efficiency of financial activities performed by the organization, for example. Organizations may utilize internal auditors to perform a similar type of audit. However, the internal auditor typically does not attest to financial reports. Rather, the focus of the internal auditor is on the controls of the
organization. External auditors may choose to place limited reliance on the work of internal auditors. Other areas commonly considered for external audits include risk assessments, quality assurance (e.g., ISO certification or recertification), compliance with Sarbanes-Oxley legislation (applicable to publicly held U.S. firms), and certifications or accreditations specific to a particular industry or product offering.

An emerging practice, which has spawned from traditional auditing and to some degree resembles evaluation, is performance auditing. In a performance audit, the traditional criteria of procedural correctness are supplemented by criteria that are more complex, including resource economy, efficiency, effectiveness, and quality (Pollitt & Summa, 1996). However, a number of arguments have been made to demonstrate the distinct differences between traditional auditing and evaluation, and the challenges, advantages, and disadvantages with integrating the two activities (e.g., Chelimsky, 1985, 1996; Divorski, 1996; Leeuw, 1996).

Level 3 Evaluation

The third level of conscious evaluation use within organizations requires the value of evaluation to be internalized as part of the organizational culture. The realization of this level includes not only levels one and two (i.e., skilled internal evaluations and use of external evaluators), but requires the full integration and acceptance of the evaluative attitude as the essence of the organization. Scriven (2008) refers to organizations that internalize this value of evaluation as evaluative organizations. The evaluative organization may be thought of as an “enhanced” learning organization—one that recognizes that learning is of no value unless it informs action. A learning organization is typically defined as one that has the capacity and processes in place to
detect and correct error or improve performance based on experience or knowledge acquired (DiBella & Nevis, 1998). Being a learning organization implies that the organization is actually learning (i.e., acquiring knowledge and implementing behavioral change) as a result of its learning processes and not simply supporting staff to attend seminars or conducting internal training for the sake of training.

Although a learning organization actively captures, transfers, and utilizes relevant knowledge to adapt to its changing environment, it does not necessarily collect information to determine the extent to which its changes are contributing or have contributed to the improvement of organizational performance. In other words, a learning organization does not specifically attempt to determine the merit, worth, or significance of a strategic initiative or business process and its contribution toward improved organizational effectiveness. To arrive at this evaluative conclusion, something more is required beyond pure learning.

In contrast to a pure learning organization, the evaluative organization fully integrates the evaluative attitude and culture into its business processes and utilizes information that has been specifically collected to determine the effectiveness of a particular process or activity and its contribution to organizational performance. What’s more, an evaluative organization recognizes that learning enables the organization to adapt to its environment and enhances organizational learning by assessing the merit of initiatives resulting from the learning process and their contributions toward improved organizational performance.

Although an evaluative organization is a learning organization, a learning organization is not always evaluative. The evaluative organization “adds value” to the learning organization concept by assessing the extent to which the knowledge acquired
by or integrated into the organization is worthwhile and used to improve organizational effectiveness. Hence, an evaluative organization is a learning organization that instinctually reflects on its actions and external environment and continuously improves as a result of those reflections. The nature of the evaluative organization is such that its evaluative actions are not simply rule-following, but a component of the organization's value system (Scriven, 2008). In other words, the *modus operandi* of an evaluative organization is its incorporation of the evaluative attitude as an implicit element of organizational culture, moving well beyond the explicit acknowledgement of a commitment to evaluation or simply doing evaluations. The evaluative organization is one that reaches beyond performance measurement and monitoring to embrace the relentless pursuit of quality and value by thinking and acting evaluatively to improve organizational performance.

**Value of Organizational Evaluation**

Beyond skilled evaluation use for quality assurance purposes, the benefits and value-added from evaluation to the organizational setting have tended to focus on learning and personnel performance initiatives. For example, Russ-Eft and Preskill (2001) reviewed various approaches evaluators have taken when assessing human resource development initiatives and highlighted the major issues associated with evaluating these activities. Other researchers have also provided insight into evaluation's use and application in organizational settings, focusing on organizational learning, training, or personnel evaluation (e.g., Brinkerhoff, 1988; Hamblin, 1974; Holton, 1996; Joint Committee on Standards for Educational Evaluation, 1988; Kaufman & Keller, 1994; and Kirkpatrick, 1959, 1994). However, the contribution of serious evaluation in
organizations goes well beyond quality assurance and human resource development. In an evaluative organization, performance improvement leads to increased competitiveness and sustainability as a result of continuous learning and reflection (Davidson, 2001). Rather than accepting current practices, procedures, and policies as forever enduring, the evaluative organization seeks out opportunities to improve its practices to maximize value to external stakeholders (e.g., consumers, funders, owners, community), as well as internal members. It leverages the evaluative culture of the organization to facilitate investment decisions, determine how best to allocate resources, direct research and development initiatives, select appropriate strategies, maximize contributions from personnel, improve training programs, and demonstrate effectiveness and accountability to its constituencies. When an organization, through its deliberate actions, incorporates the evaluative attitude into its operations as outlined above, returns are maximized and the organization thrives. In others words, the organization is effective.

Barriers to Advancing Evaluation in Organizations

Despite the lower-level uses of evaluation in organizations, there have been some valiant attempts by organizations, management consultants, and organizational researchers to advance evaluation—particularly process-focused evaluation—within organizations during the past twenty years. Initiatives such as Quality Circles, Total Quality Management, Six Sigma, and Lean Enterprise have focused organizations on process improvement, while new or updated process maturity models offer guidance for conducting process audits (e.g., Hammer, 2007; Software Engineering Institute, 2007). However, these approaches have narrowly focused on lower levels of evaluation use and have not advanced to the evaluative organization level, nor have they been expanded to
other types of evaluation (e.g., outcome or impact) that can also inform decision making, organizational planning, and strategy, among other areas. To illustrate precisely where evaluation—and specifically the evaluation of organizational effectiveness—has fallen short, it is necessary to examine some of the more prevalent barriers that have prevented evaluation from becoming an integral and necessary part of organizational life.

Value Claims in Evaluation

The limited role of evaluation use in organizations may be attributed, in part, to a misconception or a lack of awareness of the different types of value claims made in evaluation. Scriven (2007b) outlines four major types of value claims found in evaluation: (1) personal preference claims, (2) market-based claims, (3) contextual claims, and (4) essentially evaluative claims. With respect to serious evaluation, personal preference claims have limited application, except when these preferences overlap with the needs of the individual, organization, or society. In some instances, personal preferences are definitive, such as in the case of medical treatment options or when choosing a particular variety of wine for personal consumption. In these cases, the use of personal preference is a legitimate source of values. In most cases, these types of claims reflect a private state of mind and cannot be validated to any wider applicability (Scriven, 2005).

Market-based claims are those that have a standard method of verification recognized by law or other agreed-upon techniques. Illustrations of this type of value claim are found in everyday transactions in which sellers and buyers exchange objects based on a shared agreement of value. For example, the worth (i.e., market value not book value) of an organization's assets such as property, machinery, and equipment or its outputs (i.e., goods and services) can be justifiably valued for transactional purposes.
These market-based value claims are testable, recognized by external parties, and markedly objective relative to personal preference claims. Unlike personal preference claims, market-based value claims are fully appropriate for use in organizational evaluations.

Contextually evaluative claims, also referred to as implicit value claims, are those that are intrinsically descriptive, but in a particular context are imbued with evaluative significance (Scriven, 1991). For example, the description of a diamond as colorless, having five carats, internally flawless, and a standard brilliant cut suggests the quality and the worth of the diamond are exceptionally high. Other examples include the height of a basketball player, the handicap of a golfer, or the time it takes a car to get from zero to sixty. In an organizational setting, “creative” could be considered a contextually evaluative claim when referring to a graphic design boutique, just as “consistent performer” could be contextually evaluative in the case of an organization in a hypercompetitive or volatile business environment.

The fourth type of value claim is an essentially evaluative claim. Unlike personal preferences or opinions, essentially evaluative conclusions are valid and defensible and provide decision makers within the organization with valuable insight for implementing, modifying, continuing, or discontinuing a program, strategy, policy, product, or other type of evaluand. To arrive at such a claim, the synthesis of criteria of merit and factual performance data is required. It is with an essentially evaluative conclusion that organizations can realize the greatest value, advancing the purely descriptive acts of data reporting and monitoring to the well-informed evaluative conclusion that provides the insight needed by decision makers to know what is working well and what needs to be improved.
Evaluation Anxiety

Another phenomenon that has prevented the expansion of evaluation use in organizations is the commonly experienced emotional reactions to being evaluated or critiqued, such as uneasiness, uncertainty, and apprehension. This reaction is known as evaluation anxiety, which is generally defined as "the set of (primarily) affective, and also cognitive and behavioral responses that accompany concern over possible negative consequences contingent upon performance in an evaluative situation" (Donaldson, Gooler, & Scriven, 2002, p. 262). In general, the fear of being evaluated or experiencing anxiety—even high levels of anxiety—is not the major issue that is limiting the advancement of evaluation in organizations. Rather, it is the excessive or disproportionate level of anxiety. An abnormal level of anxiety leads to a number of consequences, such as a lack of cooperation among critical organizational members in evaluation activities, unconscious filtering or distortion of information reported to the evaluator, manipulation of performance data, attacks on the validity of the evaluation findings, poor utilization of evaluation results, and dissatisfaction with evaluation (Donaldson et al., 2002). In essence, excessive evaluation anxiety can undermine the quality of the evaluation and act as a major obstacle for the evaluator. Signs of excessive evaluation anxiety include exceptionally high levels of conflict between organization members and the evaluator, avoiding communication with the evaluator, resistance to providing information or access to personnel, hostile or aggressive actions toward the evaluator, delays or slow-downs in providing data or fulfillment of various requests, and dismissal of evaluation findings. A number of techniques have been suggested to alleviate or prevent excessive evaluation anxiety (e.g., Arnold, 2006; Donaldson et al., 2002). One of the more frequently mentioned approaches is to have an open and frank
discussion about the evaluation’s purpose and its potential undesirable consequences and how those risks will be managed (Donaldson, 2001). The objective is not to eliminate evaluation anxiety, but to be sensitive to its manifestations and potential sources of excessive evaluation anxiety in an effort to minimize negative consequences and maximize potential use of the evaluation results.

Performance Measurement Proliferation

As noted earlier, the growth in the field of performance measurement has increased the awareness and use of performance measurement systems in organizational settings. Its popularity and “en vogue” status has led to the development of a variety of process-oriented measurement techniques (e.g., Six Sigma) and philosophies (Lean Enterprise) that are commonly incorporated into an organization’s process management and monitoring systems. This focus on performance measurement and ongoing monitoring has narrowed the role of serious evaluation in organizations to process-oriented assessment, missing the opportunity to take advantage of a higher-level application of evaluation that supports decision making at a multiple levels. Although performance measurement is an important aspect of planning and control within the organization, it is not able to assert whether the measured performance is worthwhile or if performance is excellent or poor. To arrive at this conclusion, synthesis of performance data and criteria of merit is required. That is, a distinct and important aspect of evaluation logic is applied to arrive at an evaluative conclusion regarding the merit, worth, or significance of the evaluand. In spite of the valuable contributions of performance measurement to organizational planning and control, one of its potentially negative consequences is its inherent tendency to reward managers, employees, and staff to maintain status quo. As a
result, the active use of performance measurement systems can discourage experimentation, learning, and creativity among organizational members—including its leaders. As such, performance measurement is implicitly linked to the notion of diagnostic control systems where actual performance and planned objectives are compared in an effort to correct deviations from preset standards of performance (Henri, 2004). In traditional performance measurement systems, this discrepancy analysis is done without regard for the merit, worth, or importance of the stated objectives.

Conceptualizing Organizational Effectiveness

The highly ambiguous nature of social organizations characterized by multiple and contradictory goals and objectives, diverse constituencies, and an increasingly complex, turbulent, and interdependent external environment has stymied the development of a common conceptualization of organizational effectiveness. This has created yet another barrier to the advancement of serious evaluation in organizations. Although agreement has been reached on some aspects (e.g., the construct is multidimensional and complex, Robbins, 1990), a number of unresolved issues and noncumulative theories have hindered the development of the conceptualization of organization effectiveness. Three obstacles that have constrained the development of a common definition of organizational effectiveness are discussed below, including (1) equating efficiency with effectiveness, (2) a focus on organizational differences, and (3) the complexity of organizations.
Attempts to define organizational effectiveness by equating it with organizational efficiency are common and incorrect (Ridley & Mendoza, 1993). From an economist’s perspective, a more efficient organization is a more healthy and effective organization. However, efficiency does not ensure effectiveness. Organizational efficiency is generally understood to be a ratio that reflects the comparison of some aspect of unit performance with the costs (e.g., time, money, and space) incurred for that performance. It is often used to measure aspects of a process other than just physical output, insofar as efficiency includes a reference to the amount of resources involved. A higher ratio of energetic outputs to energetic inputs suggests a higher level of efficiency. The more efficient the operation, the less energetic return required to maintain ongoing activities. When sustained over time, the efficiency-generated surpluses result in organizational growth and survival power (Katz & Kahn, 1978). However, the efficiency ratio does not guarantee the results are of any useful size. Because of this limitation, efficiency measures are generally supplemented by other measures of organizational performance or success. Although efficiency is essential to the effective functioning of an organization, improvements in internal efficiency do not always suggest increased organizational effectiveness. For example, an organization may be highly efficient at producing vinyl (LP) records. Despite this efficiency, the refusal of the market to accept vinyl records (the organization’s output) makes the exchange with the environment impossible. The organization is unsuccessful in its ability to consummate one of the fundamental and necessary aspects of organization function—that of exchange. Therefore, the organization is ineffective in spite of its efficiency.
Organization Typology Perspective

The long-established approaches used by organizational theorists to define organizational effectiveness have emphasized different perspectives with respect to the organization type and degrees of importance of the various constituency groups comprising the organization. More specifically, definitions of the organization and organizational effectiveness have focused primarily on the dissimilarities among organizations and their constituencies, while conceding efforts to identify common denominators of organization. The focus on organization differences and the unique perspectives of particular constituent groups has led to increased fragmentation of the conceptualization of organizational effectiveness and weakened its utility.

In an effort to address the growing divergence and identify salient effectiveness criteria, several scholars (e.g., Carper & Snizek, 1980; Rich, 1992) have attempted to develop typologies of organizations or determine what major characteristics are typical of different types of organizations. However, these efforts have failed to generate a common set of criteria, presumably because no "standard" organization exists. The traditional focus on organization type has perpetuated the ambiguity surrounding organizational effectiveness and severely limited attempts to identify essential processes (i.e., functions) that are inherent across all types of organizations. By probing to identify underlying organizational functions that are shared by all organizations, it is possible to construct a universal conceptualization of organization that can be applied to evaluating organizational effectiveness regardless of the type or purpose of the organization.
The Complexity Issue

The third obstacle to conceptualizing organizational effectiveness is found with the inherent complexity of organizations. Organizational boundaries can be uncertain and fluid, goals are frequently contradictory, and competing interests are held by a variety of constituencies both internal and external to the organization. Cohen and March (1974) describe this situation as organized anarchy. This orderly chaos is characterized by organizations that function with inconsistent and ill-defined preferences, organizational members that perform their duties without understanding organizational processes, and decision makers that can change capriciously for any particular choice.

A related concept to organized anarchy used to illustrate the complex and ambiguous nature of organization is referred to as loose coupling or loosely coupled system (Orton & Weick, 1990; Weick, 2001). Organizations seen as loose coupling systems do not function with tight linkages, but as loose couplings of actors, technology, and rewards. Higher education organizations are examples of loosely coupled systems (Cameron, 1986). In these organizations, several different means can produce the same result (i.e., more than one strategy produces the same outcome), influences from the external environment are partitioned across subunits and weaken quickly on spreading, and resource acquisition in one area or subunit has no direct relationship to the outputs of another. This abbreviated list of characteristics illustrates the complexity found in organizations that are loosely coupled and supports the argument that organizations are complex and that alternative perspectives are needed to understand and evaluate organizational effectiveness.
Toward Evaluating Organizational Effectiveness

This dissertation sets forth to address the discombobulation found with current approaches to evaluating organizational effectiveness. It does so by incorporating an explicit focus on the common functionality innate to all organizations regardless of the organization size, type, structure, design, or purpose. The widely accepted open-systems perspective (Katz & Kahn, 1978) of organizations is used as the foundation for this conceptualization of the organization to explicate the similarities of organizations rather than the differences. The open-systems perspective suggests that the organization, while itself a system (i.e., situation of interest), is part of a larger system from which it acquires resources and conducts exchanges. The resources may include labor, capital, machinery, knowledge, equipment, and other inputs as required by the organization to fulfill its purpose. The inputs are coordinated, manipulated, and transformed into a specific good, service, or idea. The result of this transformation is referred to as the organization’s output. The organization’s output is exchanged with the environment in the form of a transaction involving the organization and a recipient. The value received by the organization from the recipient of the output is then used by the organization to acquire more resources (i.e., inputs). Figure 1 illustrates these five basic functions common to all organizations.
Using the input-transformation-output-exchange cycle as the fundamental premise for characterizing an organization, a common framework can be utilized for evaluating organizational effectiveness. In contrast to existing models of organizational assessment summarized in the following chapter, the framework presented herein also incorporates the logic and methodology of evaluation into the approach to address the limitations found with the prominent models in the organizational strategy and theory literature.

Thus far, I have argued that the utilization of serious evaluation in organizations has been limited by a variety of elements—not the least being the ambiguity surrounding the conceptualization of organizational effectiveness. Because of its central importance to both researchers and management practitioners, among many other organizational participants, the development of a practical tool designed to address the deficiencies and
challenges in current approaches to organizational effectiveness evaluation is of theoretical and practical value.

The aim of this dissertation is to address the lack of satisfactory approaches for assessing organizational effectiveness and present an alternative framework to evaluate organizational effectiveness. The framework provides a systematic approach toward evaluating organizational effectiveness and is a practical tool for guiding professional evaluators, managers, and organizational consultants in their evaluation efforts. Particular emphasis is placed on the importance of evaluation logic and its vital contribution to any approach for evaluating organizational effectiveness. The scope of the dissertation is limited to the evaluation of organizational effectiveness at the organizational level of analysis, using an evaluative checklist approach. This dissertation contributes to both the evaluation and management disciplines by presenting new perspectives on evaluating organizational effectiveness in hopes of augmenting existing models for a better understanding of the organizational effectiveness construct.

The first chapter of this dissertation provided an introduction to evaluation in organizations and the challenges found with respect to evaluating organizational effectiveness. The following chapter explores the elements and issues associated with evaluating organizational effectiveness and includes an overview of six prevalent approaches for evaluating organizational effectiveness. Each approach is described in terms of its development and application, followed by the identification of advantages and limitations of the specific approach. Chapter 3 explores the development and aspects of a new framework for evaluating organizational effectiveness, referred to as the Organizational Effectiveness Evaluation Checklist (OEC). Although labeled a checklist, the OEC outlines a comprehensive process approach for evaluating organizational
effectiveness. Unlike most other models of organizational effectiveness assessment, the OEC unambiguously addresses the values issue with its inclusion of universal criteria of merit. Chapter 4 discusses the methodologies used to validate the OEC framework. These included an expert panel as phase one of the investigation and a field case study applying the OEC as phase two. Findings from the two investigations are presented in Chapter 5, and Chapter 6 concludes with a discussion of the implications of the OEC for organizations and the professional practice of evaluation. Limitations of the research and opportunities for future exploration are also described in the final chapter.

Throughout this dissertation, the terms assessment and evaluation are used interchangeably and are intended to confer the same meaning. However, in some organizational settings and in organizational research, the term assessment is commonly used when referring to measurement or as a synonym for audit—not necessarily the determination of the evaluand's merit, worth, or significance. The terms effectiveness and performance are also used interchangeably, because problems related to their definition, measurement, and explanation are virtually identical (March & Sutton, 1997). It is interesting to note that prior to 1978, both terms (i.e., effectiveness and performance) were used interchangeably in the literature. Since that time, the term performance has become the more dominant concept (Shenhav, Shrum, & Alon, 1994). Of final note, the terms performance and performance measurement must not be confused, as the distinction is important. The former refers to outcome while the latter is a measurement tool.
CHAPTER II

ELEMENTS AND ISSUES IN ORGANIZATIONAL ASSESSMENT

Introduction to Organizations

An organization is a planned social unit deliberately structured for the purpose of attaining specific goals (Parsons, 1960). Since the days of Pharaohs building pyramids and Chinese emperors constructing irrigation systems, formal social groupings of humans have been constructed and reconstructed for specific purposes. Etzioni (1964) identified three broad characteristics of organizations, including (1) nonrandom and planned divisions of labor, power, and communication responsibilities to support the achievement of specific goals; (2) the presence of one or more power centers (i.e., dominant coalitions) that direct the concerted efforts of the organization toward its goals; and (3) the substitution of personnel whereby unsatisfactory members can be removed and others assigned their tasks. These characteristics further delineate Parsons' definition of an organization and introduce the concepts of power, politics, and substitutability of organization members.

There are certain types of social units, however, that are excluded from this definition of an organization, including tribes, classes, ethnic groups, and families. A primary distinguishing factor between social units that are classified as an organization and those that are not is that membership in the latter is not choice-based. In other words, an individual does not choose a particular ethnicity or the family to which they
are born. Another distinguishing factor is the lack of substitutability of personnel within social units such as families and tribes.

The term “deliberately structured” implies some type of coordinated management within the organization. More specifically, this management refers to the consciously coordinated management of the interactions among organizational members. As part of the organization structure, there are relatively identifiable boundaries that distinguish between organizational members and organizational actors, where the former include those who act legally on behalf of the organization including employees, management, boards of directors, and agents. Organization actors are those who are acted upon by the organization, including customers, vendors, and society, among other external constituencies. Organization members typically have an explicit or implicit contract with the organization to perform specific duties in exchange for compensation. In contrast, members of volunteer organizations typically contribute in exchange for intrinsic rewards, such as the satisfaction of helping others, or for social interaction—rather than monetary considerations (Robbins, 1990).

The Organic Perspective

The primary tasks confronting all organizations include environmental transactions (i.e., the supply of inputs and the exchange of quality outputs), production (i.e., the transformation of inputs), and maintenance or support of those activities and the persons performing those activities. As suggested by Katz and Kahn (1978), a basic theoretical construct for understanding organizations begins with the input, output, and functioning of the organization as a system and not necessarily with its rational purposes. This is not to say that the organization’s purpose is ignored. To the contrary,
organizational purposes provide insight into the organization’s functioning as well as its domain. However, a more pragmatic approach to understanding organizations, as outlined in the introduction, is to view the basic function of the organization as an energic input-output system with subsystems to maintain and adapt to internal and external environments acting on the system of interest. In a manufacturing organization, for example, the inputs may include raw materials, machinery, and labor. The production activities—which may include entering purchase orders from customers, scheduling routings, processing of raw materials, and machining and assembling components—are performed during the transformation process. The final product is the output. This output is exchanged for something of value (e.g., money) and the exchange provides energic input to continue the cycle. To enable this input-output cycle, individuals (i.e., organizational members) perform various activities alone or in groups. These activities, combined with the exchanges within and external to the organization, support this open-systems perspective of an organization.

The Mechanical Perspective

The organic conception of organizations summarized above may be contrasted with mechanical (Burns & Stalker, 1961) or machine theory (Worthy, 1950) conceptions that are variations of the classical bureaucratic model inspired by Weber (1947). The classical bureaucratic model and its derivatives, including the scientific management approach of Taylor (1923), emphasize process specialization of tasks, standardization, centralization of decision making, and an overriding focus on efficiency. Lean Enterprise and Six Sigma are examples of contemporary approaches that reflect the classical bureaucratic model and are used in both service and manufacturing organizations.
Modern approaches specific to manufacturing organizations include 5S, cellular manufacturing, just-in-time, statistical process control, and total productive maintenance, among others. Unlike the organic conception of organizations, bureaucratic models give limited attention to the effects on organizational participants or to changing environmental elements. Rather, these models view the organization as a rational system that works toward achieving certain objectives. Ignoring or minimizing the effects on and influence by organizational participants and the external environment limits the mechanical conception to an internal, closed-system focus and overlooks a primary task of the organization—the environmental transactions that ensure supply of material inputs and exchanges of the organization's outputs.

Based on the definitional premises of an organization (i.e., deliberate, planned, purposeful), its systemic role in modern society, and influences from within and outside of organizational boundaries, the organizational perspective used in this dissertation explicitly recognizes the importance of integrating the rational-systems perspective (Cyert & March, 1963) and the open-systems perspective (Katz & Kahn, 1978). Inasmuch, the resource acquisition (e.g., environmental transactions) element of the open-system perspective is considered one type of constraint that must be met before organizational goals can be attained.

Aspects of Organizational Effectiveness

Theoretically, the construct of organizational effectiveness lies at the heart of all organizational models. It is, in essence, the ultimate dependent variable in organizational research. However, despite the plethora of writing and research on organizational effectiveness, there remains little consensus on how to conceptualize, measure, and
explain the concept. A major challenge for organizational evaluation, therefore, is the
determination of the most useful lines for distinguishing between effective and
ineffective organizations (Cameron, 1980).

Level of Analysis

One key issue to address with respect to the study of organizational effectiveness
is level of analysis. Katz and Kahn (1978) provide a useful distinction between levels of
phenomena and levels of conceptualization in organizational analysis. The level of
phenomena refers to that which can be encountered, observed, measured, and
manipulated. The conceptual level has to do with ideas and theories about phenomena.
For example, the delivered service of a food bank is the output of a group of people, and
in that sense, it is an emergent phenomenon rather than a mere aggregation of individual
acts. However, the interrelated acts of individuals, both internal and external to the
organization, are what are directly observable in creating that delivered service. Because
the study of organizational effectiveness is at a higher level conceptually, but not
phenomenologically, Katz and Kahn (1978) argue that “the study of organizations
should take the social system level as its conceptual starting point, but that many of the
actual measures will be constructed from observations and reports of individual behavior
and attitude” (p. 13). The societal level referred to by Katz and Kahn concern the costs
and benefits associated with the organization’s continuing functions, while the
organizational level is concerned with the maximization of returns over some period.
Nature of Organization

A second issue with respect to the study of organizational effectiveness is defining the nature of the organization. In other words, defining what makes for an effective organization. A number of definitions of organizational effectiveness have been proposed, all influenced by the particular organizational perspectives held by the authors or proponents. The multiple organizational models, structures, and purposes combined with the differing theoretical and philosophical roots of organizational theorists have facilitated the conceptual disarray and methodological ambiguity surrounding the organizational effectiveness construct.

Thorndike (1949) first noted the general trend among organizational researchers to assess organizational effectiveness using a univariate model that focused on an ultimate criterion, such as profit or productivity. This approach directed attention toward organizational outcomes and viewed the organization as a rational system that enables the attainment of certain objectives. Organizational effectiveness, when described from this perspective, is the degree to which the organization attains desired end states. Despite the frequent use of univariate models, there is considerable difficulty associated with defending the use of a certain variable by itself as a comprehensive measure of organizational effectiveness when there are such a large number of variables capable of influencing effectiveness (Boswell, 1973; Steers, 1975).

Georgopolous and Tannenbaum (1957) shifted the focus from univariate to multivariate with their attempt to expand the list of criteria to include organizational flexibility and intraorganizational strain based on their empirical investigation of an industrial firm’s effectiveness. Effectiveness under this conceptualization focuses on how well an organization can achieve its objectives, given certain resource constraints,
without placing undue strain on organizational members. This systems approach does not exclude goal attainment as a measure of organizational effectiveness. Rather, it places greater emphasis on functional complementarity among the organization's parts and stresses the need for maintenance. Subsequent multivariate organizational effectiveness models incorporated one or both of these perspectives on the nature of organizations and organizational effectiveness criteria, cumulatively contributing to the multiple facets of the organizational effectiveness construct.

As outlined in the previous sections, all definitions of organizational effectiveness involve some assumptions with respect to the frame of reference (i.e., the nature of the organization) and level of analysis. There are four basic components that comprise the complex organizational system: (1) inputs, (2) transformations, (3) outputs, and (4) maintenance. The inputs may include labor, materials, physical resources, and knowledge. Labor is used to transform the inputs into a good or service using equipment, materials, knowledge, or other resources. The good or service that is created is referred to as the output. The output is exchanged with an external party (e.g., a customer or intraorganizational unit) in return for something of value. In the case of a for-profit organization, this exchange would typically include the transfer of money. In the case of a nonprofit service organization where its service is offered to recipients free of charge, the exchange might be time, effort, or something else given in return for the service provided. As a result of this transaction, a return is provided to the organization. This return allows the organization to acquire additional inputs and maintain the system.

An organization delivers value is by effectively managing and maintaining the input-transformation-output cycles while minimizing harm from its activities (Keeley, 1984). Sustainable value is conveyed in a variety of ways to organizational constituencies.
For example, continuing operations create wealth for owners, jobs for employees, exchanges with customers and vendors, and quality of life for the community. When the organization is able to sustain the input-output cycle while maximizing returns to the various constituencies, organizational functioning is considered effective. Hence, we can derive two fundamental elements applicable to all organizations: survival and maximization of returns. Neither alone is sufficient to assess organizational effectiveness. However, both are necessary. Based on these fundamental and common elements, organizational effectiveness can be defined as the extent to which the organization provides sustainable value through the purposeful transformation of inputs and exchange of outputs, while minimizing harm from its actions.

The term sustainable value in this definition refers to the value ascribed by a particular stakeholder. For example, owners may define value as wealth creation. Employees may define value as fair compensation for the work they provide. Customers may define value in terms of quality versus price. Communities may define value as a measure of the financial contributions made by the organization to the community chest, whether benevolent or through taxation. Therefore, this definition of organizational effectiveness addresses the purposeful transformation of inputs, exchange of outputs, and adaptability of an organization to deal with the external environment, all the while minimizing harm created by the actions of the organization.

Values and Organizational Effectiveness

A third issue with respect to the study of organizational effectiveness is that of values. More specifically, the sources of values used to arrive at evaluative conclusions. All organizational effectiveness evaluations consist of two components: facts and values.
However, the value-based nature of the effectiveness construct generally is ignored in most organizational effectiveness models (Zammuto, 1982).

An organization does not have a value perspective of its own. Rather, the multiple value perspectives are found outside the organization proper (Seashore, 1983). In other words, the internal and external constituencies hold various value perspectives from which they interpret organizational activities and outcomes. Because of the challenges with respect to the traditional handling of multiple value perspectives, defining the nature of the organization, and inappropriate biases resulting from poorly chosen value premises, the advancement of serious evaluation in the organizational setting has been all but abandoned. In its place, compliance audits to ensure procedures are being followed have emerged under the guise of quality assurance, financial audits, risk assessments, and other "value-free" appraisals. In some cases, pseudo-evaluations are conducted for public relations or political purposes, seeking information that would most help an organization confirm its claims of excellence and secure public support (Stufflebeam & Shinkfield, 2007). The Malcolm Baldrige Quality Award given in 1990 to Cadillac for its operational excellence is an example of a pseudo-evaluation. Despite the quality and performance problems that a serious analysis of owner surveys and road tests clearly showed, the organization was given this distinguished award based on a "sound" evaluation of the organization. (Scriven, 1991).

A number of organizational theorists have questioned the impartiality of organizational assessments, specifically referring to criteria that seemingly serve as expressions of an individual’s values and preferences for what defines an effective organization (e.g., Keeley, 1984; Steers, 1977). Some even called for a moratorium on organizational effectiveness studies due to this lack of objectivity, inability to articulate a
single model or theory of organizational effectiveness, and unanswered questions regarding a clear definition of the construct (e.g., Goodman, Atkin, & Schoorman, 1983; Hannan & Freeman, 1977). The issue, however, is not the inclusion of values in organizational studies, or even the multiple value perspectives prevalent in organizational settings. The values that are applied to descriptive facts (i.e., performance data) is what allows for a conclusion to be drawn regarding whether or not the evaluand is meritorious or valuable, whether the organization is effective or ineffective. Rather, the central issue is utilizing appropriate value premises to arrive at essentially evaluative conclusions concerning the organization's effectiveness.

Relevant values used in organizational evaluations may come from a variety of valid sources, including an understanding of the nature of the organization (i.e., criteria of merit that defines an effective organization), the organization's needs, professional standards, logical analysis, legal requirements, ethical requirements, factual and definitional premises, organizational goals or desires articulated by the dominant coalition, fidelity to or compliance with alleged specifications, cultural or traditional standards, risk, marketability, and resource economy (Scriven, 2007b). It is important to note that none of these value premises suggests the use of arbitrary or idiosyncratic personal preferences. Instead, they are based on defensible, valid sources derived from organizational needs, definitions, and ethics, among other sources.

Traditional Approaches to Assessing Organizational Effectiveness

The following section provides an overview of five classifications of models to assess organizational effectiveness. Although there are other models and approaches that could be highlighted, the following are the most widely discussed in the literature. A
sixth approach for assessing organizations, the Baldrige National Quality Award, is included in this review based on its prominence in business settings in the U.S. The models are presented in chronological order of their development. An attempt has been made to reference the original author and major proponents of the model in each case.

Goal Model

In general, the goal model defines effectiveness as a complete or at least partial realization of the organization's goals (Etzioni, 1960, 1964, 1975; Price, 1972; Bluedorn, 1980; Hall, 1980, Latham & Locke, 1991). Goals are commonly accepted as part of the organizational culture, design, and structure; and the organization's purpose is generally operationalized in the form of a specific goal or set of goals. For example, the purpose of a for-profit corporation may be to generate a return for its owners. Its specific goals may be operationalized, for example, in terms of return on equity, revenue growth, market share, productivity, and cost minimization. The specific operational goals establish the basis for developing the organization's strategy and performance measures that support its purpose. In the case of a nonprofit organization, its purpose may be to serve a particular community need. This purpose may be operationalized in terms of the number of clients served, the efficacy of its programs, efficiency of operations, and community impact. In both cases, the interwoven relationship between an organization's purpose and its goals reinforces the natural inclination and legitimacy for organizational goals to be considered a suitable criterion for evaluating organizational performance and effectiveness.

Early proponents of the goal-oriented approach to evaluating organizational effectiveness focused on outcomes or the ends resulting from the organization's
activities. The criterion model (Wallace, 1965; Dunnette, 1966; Blum & Naylor, 1968; Schmidt & Kaplan, 1971) and behavioral objectives approach (Gagne, 1962; Briggs, 1968; Glaser, 1969) both centered attention on identifying specific organizational tasks or behaviors followed by the determination of whether these sets of objective-based criteria were attained. These models were derived from individual performance measurement approaches and adapted for assessing organizational effectiveness (Campbell, 1977). Management by objectives, perhaps the ultimate goal-oriented model of effectiveness, considers the primary criterion of effectiveness to be whether or not the organization accomplished the tasks that were identified as necessary (Drucker, 1954; Odiorne, 1965; Humble, 1970). The relevant measure of effectiveness is a determination of which objectives were achieved and which were not.

The goal-achievement model presumes organizations are rational, deliberate, and goal seeking. It also incorporates a number of presuppositions, the most fundamental being that organizational goals exist. Other assumptions inherent in the model are that organizational goals are specific and measureable, realistic, operative and not simply officially stated goals, relevant to the organization’s purpose, meaningful, reflect outcomes and not means, and can be defined differently than constraints.

Constraints generally appear in organizations in the form of policy statements, directives, or decision rules that guide behavior (Pennings & Goodman, 1977). For example, achieving operating profit of a certain minimum percentage, maintaining on-time delivery at a certain level, and not doing business in countries that require political kickbacks all represent constraints on the organization. Whether achieving a particular level is a goal or a constraint depends largely on its degree of centrality to the organization’s dominant coalition (Simon, 1964). For example, some organizations may
emphasize the number of clients served as a primary focus, but are constrained by the need to maintain a minimum level of profitability. Others may consider the number of clients served a constraint for high quality service delivery. Goals become constraints when viewed as goals for the organization and held by owners, managers, or other influential groups. Instead of being internal artifacts, they become external. A second distinction between goals and constraints is that partial or full attainment of goals indicates the degree of effectiveness, whereas constraints must be satisfied as a necessary condition of organizational effectiveness. Hence, there is no upside to exceeding the referent for a constraint. It defines the minimum level of performance required for effectiveness and exceeding that level does not indicate a greater degree of effectiveness.

Because of the general acceptance and proliferation of goal and objective setting by organizations, the use of a goal achievement model to evaluate organizational effectiveness is a natural consequence. However, this familiarity comes with a number of limitations and questions about objectivity and adequacy as a measure of organizational effectiveness.

Steers (1977) advances a common argument supporting the goal model for evaluating organizational effectiveness: "The major advantage of the goal approach in evaluating effectiveness is that organizational success is measured against organizational intentions instead of against an investigator's value judgments (that is, what the organization "should" be doing)" (p. 5). Although it is a reasonable statement to suggest that the investigator's preferences should not serve as the value premise for the criteria of merit, the goal-based approach lacks impartiality as it reflects the values of some subjects (e.g., the dominant coalition) and not others. Inasmuch, goals are imbued with value judgments generally biased towards the values of organizational participants who
have the most to gain by achieving a given goal resulting in an inappropriate bias toward management and program managers.

Other challenges with respect to goal attainment as the criterion of effectiveness include the lack of specificity of goals, measurement of partial completion of goals, identification and handling of side effects, importance weighting, conflicting goals, and confusing constraints with goals, among others. Etzioni (1960) suggests that it is a basic methodological error when comparing a real state (the organization) to an ideal state (goal) as if the goal were also a real state. In a different way, Seashore (1983) supports the argument made by Etzioni, "The goal model makes eminently good sense when viewed as a model for describing the purposive forces exerted on the organizational system; it makes little or no sense when viewed as model for self-generated purposiveness within organizational systems" (p. 59). In other words, the goal model may be appropriate for describing organizational constraints, but of limited value when attempting to produce evaluative conclusions about the organization.

One variation of the goal model intended to shift the balance of power slightly in favor of consumers is goal attainment scaling (GAS) (Kiresuk & Sherman, 1968). This methodology was designed to measure partially completed goals using an individualized, scaled description of outcome achievement. The technique was originally developed for use in the mental health service industry, and was later used in other human service organizations (Kiresuk, Stelmacher, & Schultz, 1982). Its use of a rubric provides a well-defined measure of the degree of goal attainment and allows for the measurement of partially completed goals using a goal composite score. However, many of the same issues with respect to other methods for measuring goal attainment are found in the GAS technique, including bias in goal selection by individuals and managers, identifying
quantifiable goals, and assignment of numerical values for importance ratings. In addition, unrealistic scaling of goals and erroneous choice of the weighted average correlation factor used in the GAS method can inappropriately influence the outcome measure of GAS.

Systems Model

A system is broadly defined as a group of independent but interrelated elements comprising a unified whole. The prominence of systems thinking flourished following World War II, influenced heavily by von Bertalanffy's general systems theory (1950, 1956, 1962, 1968) and the concept of the open system. An open system is one that exchanges matter and energy with its environment (von Bertalanffy, 1968). Boulding (1956) defined an open system as having the property of self-maintenance, the goal of survival, and the ability to maintain its existence by adapting to the environment.

Early proponents of the systems-oriented approach to assessing organizational effectiveness included Etzioni (1960), Katz and Kahn (1966, 1978), and Yuchtman and Seashore (1967). The perspective taken by this group was that the effectiveness construct is best understood in terms of the entire organizational system and control within the environment. The successful acquisition of scarce and valued resources in an environment of continuous change suggested an effective organization. Other variants of the systems-oriented approach include an operations research focus (Ackoff & Sesieni, 1968; Ackoff, 1973) and organizational development (Beckhard, 1969; Likert, 1961, 1967). The operations-research approach focuses attention on building models of the system and subsystems to determine the optimal performance of each subsystem. Although this model relies heavily on goal definition, its emphasis on subsystems
optimization as the vehicle toward goal attainment places it in the systems model group. The organizational development (OD) approach places its emphasis on human variables within the organization, specifically on intra- and inter-group behavior. While Beckhard (1969) defined an effective organization as one that is aware of, open to, and reactive to change, Likert considered effectiveness as the extent that individuals could truly participate in the decisions that affect them.

The seminal work of Katz and Kahn (1978) applied the concept of general systems theory to organizations and suggested that the theoretical model for understanding organizations be that of an energetic input-output system where organizations are "flagrantly open systems in that the input of energies and the conversion of output into further energetic input consists of transactions between the organization and its environment" (p. 20). This perspective clearly and narrowly focuses the organization's purpose on transforming inputs into outputs and the sustainability of that process.

The systems model of organizational assessment does not necessarily disregard the organization's goals; instead, it considers them as only one of the potential criteria for measuring organizational effectiveness. While the goal-attainment approach to evaluating organizational effectiveness focuses on the outcomes or ends of an organization's activities, a systems approach attempts to include process activities or means as another dimension. A systems approach recognizes that an organization must acquire inputs and transform them into goods or services to satisfy each stakeholder's goal. This approach emphasizes the importance of creating a return, maintaining the organization as a social system, and the organization's interaction with its external
environment (Robbins, 1990). In addition, the systems approach considers the degree to which the input-output system is in balance and sustainable.

The systems approach to evaluating organizational effectiveness highlights the process an organization utilizes to achieve its purpose, not only the outcomes or impacts. This broader perspective recognizes not only what gets done, but how it gets done. In particular, the process component may also consider the legality, morality, and fidelity of the organization’s activities (i.e., transformations of inputs to outputs).

The systems model incorporates a number of presuppositions, such as (1) system boundaries can be identified, (2) there is a clear connection between the organization’s inputs and outputs, (3) the most desirable resources are available and utilized, and (4) organizational management can control the environment. With respect to the organization as the system of interest, the identification of the system is achieved by defining the boundaries for the system. The concept of system boundaries is an aspect of open-system theory and is vitally important to the evaluation of organizational effectiveness. Open-system theory emphasizes the close relationship between a situation of interest and its supporting environment and throughput (i.e., the processing of inputs to produce some output used by an outside group or system). From an open-systems perspective, the boundaries can begin to be defined by first considering the input, output, and functioning of the organization. Katz and Kahn (1978) suggest the problem of identifying boundaries is solved by “following the energetic and information transactions as they relate to the cycle of activities of input, throughput, and output” (p. 21). The effectiveness of an organization (i.e., system of interest) depends on the actual boundary judgments associated with the system. Therefore, evaluating a system of interest requires identifying the boundary judgments (Reynolds, 2007).
In organizational settings that operate in highly collaborative environments (e.g., start-up organizations in a business incubator), identifying system boundaries becomes increasingly difficult as the collaborative element creates a fluid environment where organizational boundaries are fuzzy and intercooperation loosely couples organizations for periods of time. Churchman (1971) articulated nine boundaries associated with any human activity system of interest, the first of which is defined by the activity's purpose. Using this boundary definition, it becomes increasingly manageable to address the boundary definition issue with organizations that operate in a highly collaborative or loosely coupled environment.

For most organizations, once the system has been identified it becomes feasible to demonstrate a clear connection between organization inputs and outputs. However, in highly diversified or loosely coupled organizations (e.g., university settings, conglomerates, holding companies), acquired resources have little, if any, direct connection with the organization's output. In this case, the connection between inputs and outputs and resource utilization is less clear.

One advantage of the systems model's focus on balance and sustainability is its broader perspective. Rather than simply determine whether goals were achieved, the systems approach tends to look longer term and drive efficiencies within the organization that lead to growth and survival of the existing system. Another advantage of this approach is that it considers the performance of subsystems within the organization that contribute to overall organization performance. As a result, organizational members' awareness of the interdependency of organization activities is increased.
However, the focus on maximizing system resources may constrict the organization's ability to adapt to change. The drive toward greater efficiency (i.e., maximization of return) without consideration of survival, may lead to a myopic organization whose highly efficient system becomes irrelevant to the demands of the marketplace and cannot sustain itself. As a result, an organization could be determined effective (i.e., the subsystems are optimized) even when it does not possess a competitive advantage.

Another limitation with the systems model is its neglect of primary beneficiaries or key stakeholders. Although these groups would be included when part of a subsystem, they are not considered an integral element for evaluating organizational effectiveness using the systems model. Other aspects that are neglected within the systems model include external forces beyond management's control that can impact the organization. For example, demographic or sociocultural shifts can influence demand for a particular product or service independent of how effectively the system is balanced. Externally imposed mandates or legislation can have a similar effect. These external forces are not accounted for within the systems model and can lead to inappropriate or false conclusions regarding the effectiveness of an organization.

Process Model

The third approach to assessing effectiveness builds on the foundational aspects of the goal- and systems-oriented models and places a strong emphasis on internal processes and operations of the organization. Steers (1976) suggests that the process model of effectiveness provides managers with a framework for assessing the major processes involved in effectiveness. It consists of three related components: the notion
of goal optimization, a systems perspective, and an emphasis on human behavior in organizational settings. According to Steers (1976), this multidimensional approach has several advantages over the goal model and the systems model and provides a better understanding of a highly complex topic.

In a process model, goals are used to evaluate the effectiveness of organization, but goal optimization, not goal achievement, is the focus. Goal optimization considers the desired goal based on the constraints facing the organization. However, it should not be confused with suboptimization, where less than optimal goals are intentionally pursued. Rather, goal optimization attempts to recognize environmental realities that limit or prevent certain targets from being fully realized. Goal optimization also considers the differential weighting aspect of goals based on identified constraints that can impede progress toward goal attainment.

The systems perspective in the process model of evaluating organizational effectiveness assumes an open-system perspective rather than a closed system. The open-system perspective recognizes the influence of internal and external factors on the organization. Steers (1976) identified four major categories of influences on effectiveness: “organizational characteristics, such as structure or technology; environmental characteristics, such as economic and market conditions; employee characteristics, such as level of job performance and job attachment; and managerial policies and practices” (p. 59). These four categories include both internal and external forces that influence the organization.

The third element of the process approach considers the role of individual behavior as it affects organizational performance. At this point, the model moves from an organizational-level focus to an individual-level focus. In other words, it shifts from a
macro level to a micro level of analysis. The logic behind the shift is that if employees agree with the organization's objectives, they will have stronger support for, and contribute toward, helping the organization meet those objectives. This suggests that the goals of the organization and the employee's personal goals must be in harmony. However, it does not imply shared goals, only shared activities that serve the diverse and conflicting purposes of organizational members (Keeley, 1980, 1984).

The process model does not specify criteria for effectiveness; instead, it focuses on the process of becoming effective. The model incorporates various aspects of the goal model and systems model and introduces a third aspect—the link between individual behavior and organization performance. Based on this approach, the concept of organizational effectiveness is best understood in terms of a continuous process rather than an end state (Steers, 1976).

As interest grew in process optimization, widespread development and use of process maturity models were developed. Although not specifically intended to evaluate organizational effectiveness, these models identify practices that are basic to implementing effective processes. This is followed by an evaluation of the maturity of an organization in terms of how many of the identified practices have been implemented. The Software Engineering Institute at Carnegie Mellon University, established by the U.S. Department of Defense in 1981, was the developer of the first model—the Capability Maturity Model (CMM)—to gain traction in government and commercial sectors (Software Engineering Institute, 2007). CMM was specifically designed for software engineering, with subsequent models developed for systems engineering, software acquisition, and product development.
More recently, Hammer (2007) introduced the process and enterprise maturity model (PEMM) that extended the CMM framework well beyond software engineering and acquisition concerns to organizations operating in any industry. In contrast to the CMM, the PEMM framework does not specify the ideal state of a particular process. Instead, PEMM considers five process enablers and four enterprise capabilities that facilitate high performing processes to perform well. The maturity level of the business process enabler is evaluated using a three-item response scale, and the results are used to identify where the organization should direct its attention to improve operating performance. Both the CMM and PEMM focus on the process of becoming effective rather than being effective. To that end, goal attainment and resource acquisition play a lesser role in these extensions of the process model.

Kaplan and Norton (1992, 1996) introduced the balanced scorecard, yet another extension of the process model, in the early 1990s. The balanced scorecard is a performance measurement system that provides a comprehensive view of progress towards achieving strategic goals (Kaplan & Norton, 1992). Its multiperspective framework includes both financial and nonfinancial measures, introducing a broader perspective to traditional performance measurement systems used in organizations. According to Kaplan and Norton (1996), the framework is not only a performance measurement system, but also a management tool used to design and communicate strategy within the organization, align unit and organizational goals, identify and align strategic initiatives, and focus on long-term strategy. The four perspectives of the balanced scorecard include financial, customer, processes, and learning and growth (generally focused on innovation and employee empowerment). Similar to the maturity models described earlier, the balanced scorecard attempts to focus organizational
improvement efforts on process improvement. However, it does this by defining process in more general terms that incorporate employees, customers, and internal business operations. The inclusion of organizational goals, a systems perspective, and an emphasis on employee learning and growth in the balanced scorecard framework resonates with the process model of organizational effectiveness originally advanced by Steers (1976).

The process model takes on similar assumptions as those incorporated in the systems model. That is, system boundaries can be identified, a clear connection exists between acquired resources and organizational outputs, and the most desirable resources are available and utilized, among others. It also presumes goals and constraints can be identified, measured, and properly weighted. Its inclusion of goal optimization and link to individual behavior attempts to build upon the advantages of the goal attainment model using internal processes and relevant constraints as influencers on goal-directed behavior. As the limitations found with the goal and systems models, the process model does not address external forces beyond management control that can affect the organization's performance (e.g., negative economic growth, federal mandates and regulations, sociocultural shifts). In addition, the emphasis on internal process efficiencies may distract the organization from perceiving long-term environmental changes that make the organization irrelevant or noncompetitive.

Strategic Constituencies Model

In an effort to recognize the various perspectives and values held by multiple organizational participants, the strategic constituencies model emerged. In general terms, a strategic constituency is defined as any group of individuals who have a stake in the organization. More specifically, strategic constituencies can be defined by considering
those persons who legally act on behalf of the organization (employees, advisors, agents, members of governance boards, etc.) and those who are purely external to the organization and act on their own behalf, either affecting or being affected by members’ actions (Salancik, 1984; Salancik & Pfeffer, 1978). As noted earlier, those persons who act legally on behalf of the organization are referred to as organization members and constituencies external to the organization are referred to as organization actors.

According to the strategic constituencies model, an effective organization is one that satisfices the demands of its constituencies. Various authors (e.g., Ford & Schellenberg, 1982; Gaertner & Ramnarayan, 1983; Hrebinia, 1978; Pfeffer, 1978; Pfeffer & Salancik, 1978) have wrapped their definitions of organizational effectiveness around the concept of the organization as a political entity that bargains with and satisfies multiple constituencies. Gaertner & Ramnarayan (1983) argue that “effectiveness in organizations is not a thing, or a goal, or a characteristic of organizational outlets or behaviors, but rather a state of relations within and among relevant constituencies of the organization” (p. 97). “An effective organization is one that is able to fashion accounts of its self and its activities and ways in which these constituencies find acceptable” (p. 102).

Unlike the goal, systems, and process approaches, the strategic constituencies model derives its criteria from the preferences of multiple constituencies of the organization. It likens an organization to a political arena where various participants with varying degrees of power vie for control over resources. Each of the constituencies has a unique set of values that requires consideration when assessing the organization’s effectiveness.
Zammuto (1984) outlines four sets of strategic constituency models (relativistic, power, social justice, and evolutionary) of organizational effectiveness, and suggests the fundamental disagreement among the models concerns whose preferences should be satisfied through the distribution of organizational performance outcomes. Each of these subsets of the strategic constituency model is summarized below.

The relativistic approach suggests that it is inappropriate to designate one particular constituency’s preferences as the correct one and that overall judgments of organizational effectiveness should be made by someone other than the evaluator (e.g., the consumer of the evaluation report) (Connolly, Conlon, & Deutsch, 1980; Zammuto, 1984). This approach redefines the role of the evaluator to that of a researcher who collects and reports factual data without making an evaluative claim regarding the merit, worth, or significance of the evaluand.

The power perspective seeks to appease those organizational participants who can threaten the organization’s survival. In this sense, it is based on a resource-dependence model (Pennings & Goodman, 1977; Pfeffer & Salancik, 1978). Pennings and Goodman (1977) have been explicit in advancing the concept that the dominant coalition negotiates and determines the criteria against which an organization’s performance should be assessed. The greater the power of each member in the dominant coalition, the more opportunity they have to impose their preferences for criteria on the other members of the coalition. This perspective suggests “the effective organization is one that satisfies the demands of the most powerful members of the dominant coalition so as to ensure their continued support and the survival of the organization” (Zammuto, 1984, p. 607).
At the other end of the power perspective approach within the strategic constituencies model lies the social justice perspective based on Rawls’ (1971) theory of justice. In this approach, the criteria against which organizational effectiveness is assessed is based on the perspective of the least advantaged person or group. Keeley (1978) suggests that the organizational effectiveness construct can be operationalized by applying a minimum-regret principle. In other words, a highly effective organization would be “one that minimizes the regret of its most regretful constituency” (Zammuto, 1984, p. 608).

The fourth approach, championed by Zammuto (1982), is referred to as the evolutionary perspective. While not ignoring the various constituencies of the organization, the evolutionary perspective considers constituent preferences as insufficient for evaluating organizational effectiveness due to the exclusion of the constraints placed on an organization. Zammuto argues that considering constraints is essential to define the potential limits of organization performance. In addition to recognizing the constraints placed on the organization, the evolutionary perspective also recognizes changes in preferences and constraints over time. According to Zammuto (1984), “the emphasis of the model is on the continual process of becoming effective rather than on being effective” (p. 608).

The assumptions built in to the strategic constituencies model are that survival is the ultimate criterion, organizations are political arenas, vested interests compete for control over resources, constituencies can be identified and ranked, and that the preferences of a particular constituency should be satisfied. Advantages in the model include its attempt to address the concerns of those who can most impact or ensure the survival of the organization, its recognition of multiple constituencies and their
preferences, and the consideration of distributive justice in the organization environment. However, one of its prominent limitations is its bias favoring the most powerful coalitions within the organizational environment. In addition, an organization can be found effective even when it does not possess any competitive advantage, as long as the expectations of the strategic constituencies are satisfied. Other limitations of the model concern the challenges associated with separating strategic constituencies from the larger environment and evaluating organizational effectiveness in a dynamic environment where constituency preferences can shift over time.

Competing Values Framework

The competing values framework (CVF) includes two fundamental premises: (1) there are multiple and conflicting criteria associated with assessing organizational effectiveness, and (2) multiple constituencies will give preference to certain values that differ according to their organizational perspective and the interests they represent. Quinn & Rohrbaugh (1981, 1983) initially developed the CVF to clarify the language used in defining the organizational effectiveness construct in analysis, development, and design. The methodological approach used by Quinn and Rohrbaugh (1981) in their research focused on organizational effectiveness from the perspective of organizational theorists and researchers rather than organizational participants. The first phase of the investigation utilized an expert panel to refine the list of thirty effectiveness criteria suggested by Campbell (1977) so that it contained only singular constructs pertaining to organizational units. Following this activity, a broader group of organizational theorists and researchers were asked to make judgments regarding the similarity or dissimilarity between pairs of effectiveness criteria that remained. The data were analyzed using
multidimensional scaling, and the resulting spatial model mapped seventeen effectiveness criteria into one of four organizational models (i.e., human relations, open system, goal, and internal process). The spatial model included three basic sets of competing values: flexibility-control, internal-external, and means-ends.

Quinn (1988) demonstrated that two of these sets of competing values—flexibility-control and internal-external—were sufficient to describe the organizational effectiveness construct. Figure 2 illustrates the competing values framework with the vertical axis reflecting differing organization preferences for structure ranging from flexibility to control and the horizontal axis reflecting differing organizational focus from internal, person-oriented emphasis to external, organization-oriented emphasis. The four quadrants have been named human relations, open systems, rational goal, and internal process, and are suggested to represent the underlying value orientations of most organizations (Quinn & Rohrbaugh, 1983).
The label "competing values framework" is used to illustrate the polar opposition of organizational values identified in the spatial model. It could be argued that these values lead to an inherent contradiction in terminology and result in findings that are seemingly incongruent and have no place in a convergent theory of organization. Quinn and Rohrbaugh (1983) acknowledge that the spatial model is oxymoronic and paradoxical in nature. However, "this does not require that [competing values] are empirical opposites, mutually exclusive in actual organization environments" (p. 374). In other words, it is possible for an organization to be both stable and flexible or productive with satisfied employees.
The competing values framework has been used in a number of organizational research studies to analyze organizational culture and strategy (Bluedorn & Lundgren, 1993), organizational transformation (Hooijberg & Petrock, 1993), management information systems (Cooper & Quinn, 1993), employee involvement (McGraw, 1993), and organizational communications (Rogers & Hilderbrandt, 1993). Other investigations have used the competing values framework to study organizational culture (Cameron & Freeman, 1991; DiPadova & Faerman, 1993), leadership styles and effectiveness (Quinn & Kimberly, 1984), and quality of life (Zammuto & Krakower, 1991) and to describe transitions in young organizations (Quinn & Cameron, 1983). Of the various models developed to evaluate organizational effectiveness, the competing values framework has received significant attention both in application, as noted above, and in psychometric tests that demonstrate high validity and reliability estimates (e.g., multitrait-multimethod analysis, multidimensional scaling [Quinn & Spreitzer, 1991] and structured equation modeling [Kalliath, Bluedorn, & Gillespie, 1999]).

The CVF approach, as a mapping device, is useful for organizations to visualize improvement opportunities or understand what organizational effectiveness looks like based on a stakeholder perspectives. This visualization, portrayed using an amoeba-like image, encourages and potentially stimulates discussion, but does little to offer an evaluative conclusion with respect to organizational effectiveness. Moreover, it ignores how things are accomplished (i.e., the process), which necessarily leads to an avoidance of the ethicality and legality of the organization's operation. Another deficiency with the framework is its lack of attention to output quality. It subsumes this criterion in each of the eight criteria (Quinn & Rohrbaugh, 1983). Other deficiencies found with the CVF
include the omission of fiscal viability (i.e., cost-effectiveness) and environmental and social impacts from organizational actions among organizational effectiveness criteria.

Baldrige National Quality Program (BNQP)

The Baldrige National Quality Program is a U.S.-based initiative that provides a framework for assessing organizations using the Baldrige Criteria for Performance Excellence. Because of its prominence in U.S. business and industry, it is included in this review of organizational effectiveness evaluation models.

The Malcolm Baldrige National Quality Award was created by Public Law 100-107 and signed into law on August 20, 1987. The award program created a public-private partnership that coupled private-sector quality audits with national awards. The rationale behind the enactment of the legislation was that a national quality program would help to improve quality and productivity of U.S. companies (Malcolm Baldrige National Quality Improvement Act, 1987). The BNQP is managed by the National Institute of Standards and Technology (NIST), an agency of the U.S. Department of Commerce. The American Society for Quality (ASQ) manages the Malcolm Baldrige National Quality Award under contract to NIST. The BNQP’s Board of Examiners is responsible for evaluating the award applications, making site visits, and preparing feedback reports to the applicants. The board of examiners consists of approximately 570 individuals in 2008 and come from U.S. businesses and education, health care, and nonprofit organizations. The national award program has wide recognition in the U.S. and has been used as a model for a number of countries who have created their own Baldrige-type program (Baldrige National Quality Program, 2008).
The criteria used by the BNQP to assess organizational performance have evolved from having a specific focus on manufacturing quality to a focus on overall organizational competitiveness (Baldrige National Quality Program, 2008). The criteria are reviewed annually by a board of overseers, and recommendations to update the criteria or other elements of the BNQP are made to the U.S. Secretary of Commerce and the director of NIST. The 2008 criteria consider process and results, and its seven categories include (1) leadership; (2) strategic planning; (3) customer and market focus; (4) measurement, analysis, and knowledge management; (5) workforce focus; (6) process management; and (7) results. Each of the process-related criteria includes two subcriteria. The results criterion includes six subcriteria, four of which were assessed as part of the process-related criteria.

The scoring guidelines outlined in the BNQP (2008) Criteria of Performance Excellence provide the weighting given to each criterion and subcriterion, but do not provide an explanation of how the importance weightings were developed. According to BNQP (2008), the applicant organization is to indicate the areas of greatest importance with respect to key business factors in its written organizational profile and in other areas of the application. However, the scoring guidelines do not allow for the criteria’s importance weightings (i.e., maximum point values) to be adjusted to meet the specific contextual issues impacting the organization’s performance. The 2008 Criteria of Performance Excellence point values for each criterion and subcriterion are presented in Table 1.
Table 1

Baldrige National Quality Program Point Values for 2008 Criteria and Subcriteria

<table>
<thead>
<tr>
<th>Criteria and Subcriteria</th>
<th>Maximum Point Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Leadership</td>
<td>120</td>
</tr>
<tr>
<td>Senior leadership</td>
<td>70</td>
</tr>
<tr>
<td>Governance and social responsibilities</td>
<td>50</td>
</tr>
<tr>
<td>Strategic planning</td>
<td>85</td>
</tr>
<tr>
<td>Strategy development</td>
<td>40</td>
</tr>
<tr>
<td>Strategy deployment</td>
<td>45</td>
</tr>
<tr>
<td>Customer and market focus</td>
<td>85</td>
</tr>
<tr>
<td>Customer and market knowledge</td>
<td>40</td>
</tr>
<tr>
<td>Customer relationships and satisfaction</td>
<td>45</td>
</tr>
<tr>
<td>Measurement, analysis, and knowledge management</td>
<td>90</td>
</tr>
<tr>
<td>Measurement, analysis, and improvement of organizational</td>
<td></td>
</tr>
<tr>
<td>performance</td>
<td>45</td>
</tr>
<tr>
<td>Management of information, information technology, and</td>
<td></td>
</tr>
<tr>
<td>knowledge</td>
<td>45</td>
</tr>
<tr>
<td>Workforce focus</td>
<td>85</td>
</tr>
<tr>
<td>Workforce engagement</td>
<td>45</td>
</tr>
<tr>
<td>Workforce environment</td>
<td>40</td>
</tr>
<tr>
<td>Process management</td>
<td>85</td>
</tr>
<tr>
<td>Work systems design</td>
<td>35</td>
</tr>
<tr>
<td>Work process management and improvement</td>
<td>50</td>
</tr>
<tr>
<td>Results</td>
<td>450</td>
</tr>
<tr>
<td>Product and service outcomes</td>
<td>100</td>
</tr>
<tr>
<td>Customer-focused outcomes</td>
<td>70</td>
</tr>
<tr>
<td>Financial and market outcomes</td>
<td>70</td>
</tr>
<tr>
<td>Workforce-focused outcomes</td>
<td>70</td>
</tr>
<tr>
<td>Process effectiveness outcomes</td>
<td>70</td>
</tr>
<tr>
<td>Leadership outcomes</td>
<td>70</td>
</tr>
<tr>
<td>TOTAL</td>
<td>1,000</td>
</tr>
</tbody>
</table>

Note: The above criteria and scores are for assessing manufacturing, service, and small business organizations.
Advantages of this assessment approach include the comprehensive documentation available from BNQP, specific criteria of merit for assessing the organization, a strong emphasis on organizational outcomes, and specific guidelines for how to score the assessment. Comparisons to other organizations, both competitors and pan-industry, are also an integral part of the assessment process.

The primary limitation of this approach to organizational effectiveness evaluation is that the BNQP does not go much beyond assessing the quality of the operation. In other words, it is not a specific approach for evaluating organizational effectiveness. Rather, it emphasizes the quality of operations and how that impacts organizational performance and market success. Moreover, there are two gaps in the coverage that seem to limit its applicability to evaluating organizational effectiveness: ethical standards are not considered unless they are listed among the company’s own standards; and product and service outcome quality is market success (i.e., customer satisfaction), which may not reflect actual organizational excellence. Other deficiencies include its inability to detect side effects of organizational activities, the inability to modify the criteria importance weightings, and the cost to participate in competing for the Malcolm Baldrige Quality Award. This approach incurs pecuniary, time, and opportunity costs resulting from the assessment being primarily a written self-assessment conducted by internal resources. According to BNQP (2008), pecuniary costs for service and manufacturing organizations applying for the award is $6,150. A site visit for award finalists ranges from $20,000 to $35,000. Nonprofit fees are approximately one-half these amounts. The time required to conduct the self-assessment, the cross-organizational involvement of personnel, and site visit management can be extensive and often dwarf the direct costs associated with the application (A. VanArsdale, personal
communication, March 12, 2008). Because of the status associated with being a Baldrige award winner, the incentive for an applicant to manipulate the system may be stronger than when an award is not in play.

Summary

The pervasiveness of organizations in contemporary society suggests that the pursuit and study of organizational effectiveness will continue to be of value to both organizational participants and society. Each of the models summarized in this chapter has contributed to a deeper understanding of the organizational effectiveness construct through its distinct conceptualization of organizational effectiveness. The focus of each model on differences among organizations, however, has led to noncumulative efforts toward a common framework, and none of the models has emerged as a dominant approach to evaluating organizational effectiveness.

Contributions from the evaluation field to organizational theory and the study of organizational effectiveness have been sparse with few exceptions. The challenging issues identified by organizational and management theorists are shared with a number of domains in social and behavioral science, including evaluation. Some 30 years ago, Campbell (1977) warned against taking too parochial a view in the organizational theory field, suggesting those who are serious about the study organizational effectiveness consult fields such as evaluation research, among others, for advances that address the common issues identified in assessing organizational effectiveness. Unfortunately, Campbell’s advice has gone largely unheeded as evidenced by the lack of reference and usage of evaluation logic and methodology in organizational assessments. The following chapter presents a new framework for evaluating organizational effectiveness with an
explicit focus on evaluation logic and its contributions to advancing the evaluation of organizational effectiveness.
CHAPTER III

A CHECKLIST APPROACH TO EVALUATING ORGANIZATIONAL EFFECTIVENESS

Organization survival and long-run viability require effectiveness. Yet the models offered to assess organizational effectiveness have been plagued with numerous challenges to their utility outside of the specific paradigms used to develop the models. For example, issues concerning impartiality, narrow or inappropriate value premises, rationale for selecting the criteria of merit, differential weighting of criteria, balancing short-run with long-run interests, contextual sensitivity, synthesis of multiple evaluative findings, generalizability, and a lack of ethical consideration are ubiquitous. Considering the importance and centrality of organizations to modern society, developing a framework that surmounts these problems and advances the understanding of organizations and organizational effectiveness is a worthy effort. The checklist approach to evaluating organization effectiveness presented in this chapter is intended to meet this challenge. It addresses the deficiencies in existing organizational assessment models by focusing on the similarities in organizational function (i.e., input-output cycle) and the ultimate criteria of survival and maximization of returns as outlined in Chapter 1. It also relies heavily on evaluation logic and methodology to address the more prominent issues in assessing organizational effectiveness. Following an overview of the purpose of evaluation checklists, the new framework is presented.
The Role of Evaluation Checklists

A checklist is a mnemonic device that consists of a list of activities, items, features, properties, aspects, components, and criteria used to perform a certain task (Scriven, 2007a). When used in evaluation, checklists provide guidance for the collection of relevant evidence used to determine the merit, worth, or significance of an evaluand. The inherently systematic process found in the use of a checklist makes it highly relevant and useful for evaluative purposes. Scriven (2007a) describes seven general conclusions about the value of checklists, paraphrased as follows:

1. Checklists reduce the chances of forgetting to check something important; they reduce errors of omission directly and errors of commission indirectly.

2. Checklists, in general, are easier for the layperson to understand and validate than most theories or statistical analyses.

3. Checklists reduce the influence of the halo effect by forcing the evaluator to consider each relevant dimension of possible merit.

4. Criteria of merit checklists mitigate the tendency to see what one wants to see by forcing a separate judgment on each dimension and a conclusion based on these judgments.

5. The use of a valid criteria of merit checklist eliminates the problem of double counting.

6. Checklists often incorporate a vast amount of specific knowledge about the particular evaluand in a parsimonious manner. Hence, checklists are a form of knowledge about a domain, organized to facilitate the evaluation in an expeditious manner.

7. In general, evaluative checklists can be developed more easily than theories of the evaluand; consequently, we often can evaluate where we cannot explain. Therefore, checklists can contribute substantially to the improvement of validity, reliability, and credibility of an evaluation; and the useful knowledge about a domain.

Criteria of merit checklists, referred to in items four and five above, are of particular importance for evaluation purposes. In contrast to nonevaluative types of
checklists (e.g., a flowchart showing a sequence of events or a set of assembly instructions for an industrial product), a criteria of merit checklist features explicit criteria used to determine the merit, worth, or significance of the evaluand.

Criteria of merit checklists can be used for scoring via the numerical weight-and-sum method (Scriven, 1991). The Baldrige National Quality Award, reviewed in the previous chapter, uses a numerical weight-and-sum methodology based on its criteria of merit list. The general approach when using the numerical weight-and-sum procedure is commonsensical—weight the criteria, assign a performance score to each criterion, and sum the product of these two numbers for each criterion. The simple algorithm leads to an overall score that can be used to rank multiple evaluands. However, the approach can often lead to erroneous conclusions due to a number of logical flaws associated with the approach—including the assumption that a single scale can be used for weights, performances, and number of criteria (Scriven, 1994). Other shortcomings of the traditional numerical weight-and-sum approach include the exclusion of nonquantifiable data, no consideration for minimal levels of acceptable performance (i.e., no bars), haphazard assignment of weights, and swamping (i.e., a large number of trivial criteria “swamp” critically important criteria and precipitate invalid scores). Although modifications can be made to the traditional numerical weight-and-sum approach in response to most of these issues, an alternative scoring procedure, referred to as qualitative weight-and-sum, is a practical and oftentimes more appropriate choice (Scriven, 1991, 1994; see also Davidson, 2005).
Organizational Effectiveness Evaluation Checklist

The Organizational Effectiveness Evaluation Checklist (OEC) is a tool for professional evaluators, organizational consultants, and management practitioners to use when conducting an evaluation of organizational effectiveness. It outlines a systematic process to assess organizational effectiveness that addresses many of the deficiencies found in existing models for evaluating organizational effectiveness. The framework is intended to offer a practical method for evaluating organizational effectiveness that provides direction for the evaluator and value to the organization.

The approach used to guide the development of the OEC was based on Stufflebeam’s (2000) *Guidelines for Developing Evaluation Checklists: The Checklist Development Checklist (CDC)*. The twelve major steps of the CDC include:

1. Focus the checklist task.
2. Make a candidate list of checkpoints.
3. Classify and sort the checkpoints.
4. Define and flesh out the categories.
5. Determine the order of categories.
6. Obtain initial reviews of the checklist.
7. Revise the checklist content.
8. Delineate and format the checklist to serve the intended uses.
9. Evaluate the checklist.
10. Finalize the checklist.
11. Apply and disseminate the checklist.
12. Periodically review and revise the checklist.
Special emphasis was given to the first step (i.e., focus the checklist task) in an effort to establish a firm foundation for the OEC. In particular, a comprehensive literature review and discussions with organizational consultants, business professionals, and other relevant experts provided the basis for clarifying and justifying the criteria included in the checklist. As revisions and clarifications were made to the OEC, it was useful to work through the CDC in an iterative fashion as new insight and perspectives were gained. In fact, various aspects of steps five through nine were repeated numerous times up until the OEC was applied in a field study.

The OEC is iterative and weakly sequential. The iterative nature suggests that the user will benefit from going through the checklist several times as new information is discovered or problems identified that may require modifications to reach an appropriate appraisal for each checkpoint. For example, the identification of performance-level needs is a checkpoint under Step 2 (conduct a performance needs assessment). The performance-level needs are used to generate contextual criteria (Step 3) for which performance data are collected (Step 4). While collecting performance data, previously unrecognized environmental conditions (i.e., increasing legislative activity), may be documented and justify the inclusion of additional contextual criteria based on a newly identified performance need. These developments should be anticipated, even sought out, to ensure due consideration is given to each conclusion made during the evaluation process.

A weakly sequential checklist, in contrast with a strongly sequential checklist, is one in which the order or sequence of the checkpoints is of some importance for efficiency or psychological reasons rather than of logical or physical necessity (Scriven, 2007a). With respect to the OEC, the sequence of the checkpoints is of logistical
importance (i.e., criteria must be defined prior to collecting performance data) and ordered for efficiency. Step 1, for example, requires the evaluator to establish the boundaries of the evaluation. As a final checkpoint of this initial step, an evaluability assessment is to be performed. This activity allows for the determination of the appropriateness of conducting an evaluation. In other words, it is used to determine if the evaluand is “ready” to be evaluated. If it is determined that the evaluand is not evaluable with a reasonable degree of confidence with the available resources, the evaluation is discontinued or redirected.

The OEC is designed for use in business and industry. Nonetheless, the explicit focus on common organizational activities (input, output, transformation, and exchange) allows the checklist to be used in virtually all organizational settings regardless of the organization’s type, purpose, structure, or environment. Although the basic framework can be applied to nearly all organizations, the specific measures used may differ according to the organization type (e.g., for-profit or nonprofit), purpose, and other contextual or environmental matters. The contingency aspect of the OEC approach permits this flexibility to account for changes in environmental conditions and shifts in organizational direction. Other characteristics of the OEC include its temporal relevance (i.e., balancing short-run considerations with long-run interests), ability to deal with conflicting criteria, practical versus theoretical significance, modest usage of organizational goals, and generalizability of the approach.

The last two characteristics identified above require further explanation as organizational goals are commonly used as the criteria for assessing organizational effectiveness, and the inability to generalize is frequently used as a basis for arguing that a single approach to assess different types of organizations is not appropriate. With
respect to organizational goals and their relevance to evaluating organizational effectiveness, the OEC does not emphasize organizational goals as criteria of merit. Rather, it relies on the notion that the definitional premise of organizational effectiveness is the extent to which the organization provides sustainable value through the purposeful transformation of inputs and exchange of outputs. Goals are useful for planning and controlling organizational activities, as well as for directing human behavior, but are not the “be all end all” criteria for evaluating organizational effectiveness. Organizational goals are considered a value source (i.e., criteria) when they manifest themselves as the result of a needs assessment or are unambiguously defined and directly related to the actual operations of the organization (e.g., operative goals). The modest use of organizational goals as contextual criteria of merit provides flexibility to address the uniqueness of the organization and shifts attention from what an organization should be doing to what an organization is actually trying to do.

The generalizability issue concerns questions of external validity; that is, the extent to which the model is considered valid in a variety of organizational settings. If the model and evaluative criteria are found to be generalizable, they can be applied to all or nearly all organizations. When limited to the traditional conceptualizations of organizational effectiveness that emphasize differences among organizations, the challenges against generalizability cannot be easily defended. However, as has been argued elsewhere in this dissertation, when the conceptualization of organizational effectiveness is at the organizational level of analysis and specifically addresses the common functions of all organizations, it is possible to derive a cross-organizational conceptual premise on which to develop universal criteria of merit for evaluating organizational effectiveness.
Regardless of the complexity of organizations or the variety of their purposes and structures, it stands to reason that the common input-output cycle of all organizations and the desire to sustain and maximize returns from this cycle allow the evaluative criteria to be definitionally connected to organizational effectiveness. That is to say, universal criteria of merit can be identified and applied in evaluations of organizational effectiveness. Note that this argument for universal criteria of merit does not suggest that performance standards and measures will be universal, nor does it suggest that context-specific criteria generated from a needs assessment cannot be included. On the contrary, contextual criteria increase the checklist’s validity and applicability, and the standards and measures will vary according to the unique organizational characteristics and environment. Therefore, generalizability of the approach can be achieved when care is taken to focus on the input-output cycle common to all organizations and the resultant definition of an effective organization is used as the premise for the universal criteria of merit, while organization-specific matters are reflected in contextual criteria of merit derived from the performance needs assessment.

Outline of the OEC

1. Establish the boundaries of the evaluation.
   
   A. Identify the evaluation client, primary liaison, and power brokers.
   
   B. Clarify the organizational domain to be evaluated.
   
   C. Clarify why the evaluation is being requested.
   
   D. Clarify the timeframe to be employed.
   
   E. Clarify the resources available for the evaluation.
F. Identify the primary beneficiaries and organizational participants.

G. Conduct an evaluability assessment.

2. Conduct a performance needs assessment.
   A. Clarify the purpose of the organization.
   B. Assess internal knowledge needs.
   C. Scan the external environment.
   D. Conduct a strength, weakness, opportunity, and threat (SWOT) analysis.
   E. Identify the performance-level needs of the organization.

3. Define the criteria to be used for the evaluation.
   A. Review the universal criteria of merit for organizational effectiveness.
   B. Add contextual criteria identified in the performance needs assessment.
   C. Determine the importance ratings for each criterion.
   D. Identify observable and concrete measures for each criterion.
   E. Identify performance standards for each criterion.
   F. Create performance matrices for each criterion.

4. Plan and implement the evaluation.
   A. Identify data sources.
   B. Identify data collection methods.
   C. Collect and analyze data.

5. Synthesize performance data with values.
   A. Create a performance profile for each criterion.
   B. Create a profile of organizational effectiveness.
   C. Identify organizational strengths and weaknesses.
6. Communicate and report evaluation activities.

A. Distribute regular communications about the evaluation progress.
B. Deliver a draft written report to client for review and comment.
C. Edit report to include points of clarification or reaction statements.
D. Present written and oral reports to client.
E. Provide follow-on support as requested by client.

Process Overview

The OEC has six steps: (1) establish the boundaries of the evaluation; (2) conduct a performance needs assessment; (3) define the criteria of merit; (4) plan and implement the evaluation; (5) synthesize performance data with values; and (6) communicate and report evaluation findings. Within each step are a series of checkpoints intended to guide the evaluator. Embedded within this framework are the universal criteria of merit for evaluating organizational effectiveness. While a criteria of merit checklist would have qualified as a standalone evaluative checklist of practical use, the novelty of organizational evaluation and limited available guidance on how to conduct a serious evaluation in an organizational setting justified the development of a larger framework to facilitate evaluation efforts. The remainder of the chapter delineates each step and checkpoint and includes a rationale for the universal criteria of merit embedded in the OEC.

Step 1: Establish the Boundaries of the Evaluation

Establishing the boundary of the evaluation explicitly defines what is and is not included in the evaluation. The complexity of an organization, its multiple constituencies
and perspectives, and the open-system environment require that the effectiveness construct be bounded at the outset of the evaluation (Cameron & Whetten, 1983). The checkpoints included in this step address specific issues that allow the scope of the assessment to be defined and appropriately circumscribed.

A. Identify the evaluation client, primary liaison, and power brokers. The evaluation client is generally the person who officially requests and authorizes payment for the evaluation. It is also the person to whom the evaluator should report. However, in some cases, the client may consist of a committee or governance board. In such cases, it is important that a chairperson or group leader be identified to promote direct communication and accountability. The primary liaison is generally the primary source of information with respect to the evaluation's purpose and timeframe and is the designated interface between the evaluator and the organization. This individual is also the “go-to” person when issues arise related to data or personnel accessibility. In some organizations, an evaluation may be contracted to establish a case for or against power brokers in the environment. Because of this situation, it is important to identify the person or persons in a position of power to leverage or utilize the evaluative conclusions.

B. Clarify the organizational domain to be evaluated. The organizational domain refers to specific attributes associated with the organization. The domain is typically circumscribed by the constituencies served, technology employed, and outputs (i.e., goods or services) produced (Meyer, 1975). The organization's primary activities, competencies, and external forces influence and shape the organizational domain. In other words, the organizational domain is defined by the operating environment with special focus on identifying customers or clients served (e.g., the target markets), products and services offered, and competitive advantages of the organization. In a
loosely coupled organization (e.g., a holding company or university), there may be several domains in which the organization competes or specializes. The same condition can be found in multidivisional organizations, where the domain under investigation may be a strategic business unit or a stand-alone operating division within a business unit. Identifying the organizational domain to be evaluated is an important consideration to focus the evaluation, as well as to ensure the evaluation is directed toward the proper domain to avoid misleading evaluative conclusions about the organization's effectiveness.

C. Clarify why the evaluation is being requested. The purpose of an organizational effectiveness evaluation affects the type of data required, data sources, degree of evaluation anxiety present or that could develop, amount of cooperation or collaboration required of the organizational participants, and overall assessment strategy, among other factors (Brewer, 1983; Flooden & Weiner, 1978). A formative evaluation to identify areas for improvement will be quite different from an evaluation to determine which division to close or agency to fund. Some reasons why an evaluation of organizational effectiveness may be requested are to identify areas for improvement in a particular business unit; to facilitate prioritization of strategic initiatives regarding organizational performance; to assist decision making regarding the allocation of resources; to improve the value the organization delivers to its primary beneficiaries; and to strengthen an organization's competitive position in the marketplace.

D. Clarify the timeframe to be employed. Judgments of effectiveness are always made with some timeframe in mind. It is important to clarify over which period effectiveness is to be evaluated. Long-term effectiveness may look much different when short-term effectiveness criteria (e.g., monthly profitability) are used. It is most practical to consider
a timeframe of one year to five years. Anything less than one year may not fully reflect the contribution of various strategies and initiatives that require some period of maturation to show effect. Periods longer than five years may become irrelevant due to the complex and generally turbulent environment in which most organizations operate. For example, when the organization competes in a high-tech environment with product lifecycles of six months from introduction to obsolescence, a timeframe of more than two years becomes immaterial to the assessment of the organization's ability to maximize returns and survive in the rapidly changing environment. For most organizations, a three-year perspective is long enough to allow strategic initiatives to take effect and trends to be identified, yet short enough to maintain relevance to the operating environment. When conducting an ascriptive evaluation of organizational effectiveness, the timeframes noted above do not apply.

E. Clarify the resources available for the evaluation. The common resources required for all evaluations include money, people, and time. It is important to clarify the financial resources (i.e., budget) available to conduct the evaluation, who will be available as key informants, who will facilitate the data collection, who will approve the evaluator's request for data, and the time available for the evaluation. When an external evaluator conducts the evaluation, these items are generally clarified in the evaluation proposal or soon after its acceptance. However, it is also important for internal evaluators to clarify the resource issue upfront to avoid running out of time or funding and ensure access to personnel and data. Moreover, it is important to identify which organizational resources are available for the evaluator's use, who will provide permission to collect data, and who will confirm the authorization for doing so.
When members of the organization play a role in data collection or substantively participate in other aspects of the evaluation, the evaluator may have to factor in a longer timeframe to complete some activities since these individuals (i.e., collaboration members) may be required to perform the new evaluation assignments in addition to their current duties. What's more, collaboration members may require special training to perform the tasks adequately. A systematic and thoughtful approach should be used for the identification of collaboration members, scheduling critical evaluation activities, clarifying the roles of the evaluator and collaboration members, and other aspects associated with performing a collaborative evaluation (Rodriguez-Campos, 2005).

F. Identify the primary beneficiaries and organizational participants. Organizations operate in dynamic and complex environments that require the consideration of contextual criteria of merit for the specific organization being evaluated. Based on this requirement, the dominant coalition or primary beneficiaries are engaged to facilitate the performance needs assessment (Step 2), identify contextual evaluative criteria, and determine importance weightings of the criteria of merit (Step 3).

The identification of the primary beneficiaries serves an important function with respect to contextual criteria of merit. Although there are a number of persons and groups that benefit directly and indirectly as a result of organizational activities, the primary beneficiaries are stakeholders that are uniquely served by the outcomes of the organization’s activities. In the case of most businesses, the organization was created to generate wealth for its owners. Hence, owners would be the primary beneficiary. For a nonprofit organization (i.e., churches, schools, hospitals), the primary beneficiaries are the intended recipients of the service. For commonwealth or government organizations, the public is the primary beneficiary.
In organizations where the primary beneficiaries are not in a position to be knowledgeable of organizational functions and constraints (e.g., students in an elementary school), the stakeholder focus shifts from primary beneficiaries to the dominant coalition. The dominant coalition, sometimes referred to as the power center, includes organizational participants (e.g., parents of elementary school students) that direct the organization in its concerted efforts to achieve specific objectives. In essence, this group’s support or lack thereof impacts the survival of the organization (Pennings & Goodman, 1977; Zammuto, 1984).

Beyond the primary beneficiaries and dominant coalition, a broader group of organizational participants should be identified to aid in identifying side impacts, both positive and negative, that affect nonintended audiences. Organizational participants can be considered from two perspectives: (1) those persons who act on behalf of the organization and (2) those who are external to the organization acting on their own behalf and either affect members’ actions or are affected by them (Pfeffer & Salancik, 1978; Salancik, 1984). Those persons who act legally on behalf of the organization are referred to as organizational members and include employees, management, advisors, agents, and members of governance boards, among others. The organizational participants external to the organization—for example, shareholders, customers, vendors, and government agencies, among others—are referred to as organizational actors. The identification of organizational participants leads to the identification of those persons who are affected by organizational activities and have a stake in the organization’s survival and maximization of returns.

G. Conduct an evaluability assessment. The information collected to this point allows the evaluator to perform an evaluability assessment prior to committing significant
resources to an evaluand. Evaluability assessment is the determination of the appropriateness of conducting an evaluation. It is used to determine if the evaluand is "ready" to be evaluated based on the existence of goals, resources, data accessibility, and how the evaluation is intended to be used. It is useful in situations where goals of the organization are known, but the measures of the goals are not yet defined. What's more, the evaluability assessment allows for a preliminary review of the appropriateness of the organization's goals. Although the evaluator may not have specific knowledge about the organization's resources or capabilities at this point in the evaluation, the organization's goals can be assessed in terms of the degree to which they are specific, measurable, attainable, realistic, and time-based. In addition, official goals can be restated as operational goals and the distinction between goals and constraints can be made clear.

Conventional evaluability assessment considers four questions (Wholey, 1994): (1) What are the goals of the organization? (2) Are the goals plausible? (3) What measures are needed and which are available? (4) How will the evaluation be utilized? Scriven (2007b) outlines the general applicability of the evaluability assessment as follows:

Since everything is evaluable, to some extent in some contexts, the issue of evaluability is a matter of degree, resources, and circumstance, not of absolute possibility. Hence, while everything is evaluable, not everything is evaluable to a reasonable degree of confidence, with the available resources, in every context, for example, the atomic power plant program for Iran in April 2006, when access was denied to the U.N. inspectors. (p. 18)
It is recommended to make a quick trial run through the OEC prior to committing to the evaluation to ensure an evaluation of the organization's effectiveness is appropriate based on the questions to be answered and the data that will need to be obtained to make evaluative conclusions.

*Step 2: Conduct a Performance Needs Assessment*

The performance needs assessment provides background information about the organization and its operating environment. The systematic approach results in the identification of performance-level needs to provide the evaluator with insight to guide the evaluation, set priorities, and develop contextual criteria to supplement the universal criteria of merit. In contrast to wants or ideals, needs are things that are essential for organizations to exist and perform satisfactorily in a given context. As such, a performance-level need, sometimes referred to as a fundamental need, is something without which dysfunction occurs. This perspective goes beyond the discrepancy definition where needs are defined as the gap between the actual and the ideal, as it considers a greater distinction between different types of needs (e.g., met, unmet, conscious, and unconscious needs) (Scriven, 1990; see also Davidson, 2005; Thomson, 1987).

The performance needs assessment explores and defines the current state of the organization from internal and external perspectives. It considers the organization’s structure, strengths and weaknesses, opportunities available, and constraints that limit or threaten the organization’s survival or its maximization of return. In addition to providing the required information for identification of performance-level needs, the benefits of conducting a needs assessment include building relationships among those
who have a stake in the situation, clarifying problems or opportunities, providing baseline performance data, and setting priorities for decision making.

A. Clarify the purpose of the organization. This checkpoint is used to identify contextual aspects that may be unique to the organization or require additional inquiry. The purpose of the organization is its \textit{raison d'être}—its reason for existence. The purpose is usually, although not always, found in the organization's mission and vision statements. In cases where a written mission or vision statement does not exist, the organization's purpose can be identified by interviewing senior management and reviewing strategic plans or other organization-specific documents such as departmental plans and objectives, as well as the organization's website.

The mission of an organization is generally action-oriented and promotes the official goals of the organization. Official goals are also referred to as public goals and may not be the actual (operative) goals of the organization. Care should be taken to distinguish between official goals and operative goals, the latter being more relevant when assessing organizational effectiveness. The vision statement provides a sense of direction, is aspirational, and indicates where the organization wants to be in the future. It commonly includes the organization's purpose as well as its values. A vision statement expresses the end, while the mission expresses the means toward reaching the ends. In organizations where divisional-level or program-level goals do not align with those of the larger organization, serious attention should be given to clarifying the "true" purpose of the organization to avoid misleading evaluative conclusions about the organization's effectiveness.

B. Assess internal knowledge needs. This checkpoint is intended to provide a "pulse check" to get an indication of the health of the organization in terms of knowledge
management practices. The output of this checkpoint is documentation of (1) the existing knowledge management practices; (2) knowledge that is required, whether in existence or not, to maximize returns and support long-term sustainability of the organization; and (3) specific knowledge-management needs based on the discrepancy between the actual and required states. As a component of the performance needs assessment, this checkpoint provides another perspective on the organization's needs, with particular emphasis on knowledge management.

The first element of the knowledge needs assessment attempts to define the deliberate knowledge management efforts that are in place. Evidence for these efforts can be found in process flow diagrams, standard operating procedures, training programs, role-playing (i.e., rehearsal) activities, decision-making drills, and collaboration systems, among other sources (Ward, 2008). The second element considers what information is required to make decisions. This information may already exist as a part of the organization's current knowledge management practices or it may be desired to facilitate decision making activities. Asking questions such as, “What information would make your job easier?” or “How does the organization solve problems?” may reveal unmet knowledge needs within the organization. The third element considers the discrepancies between the current knowledge management practices and those required, but currently not available, to avoid dysfunction and enable the organization to maximize returns and sustain itself. Based on the findings of the knowledge needs assessment, contextual criteria of merit may be developed for inclusion in the evaluation.

C. Scan the external environment. The open-system nature of the organization recognizes that external factors affect organizational performance. The purpose of this checkpoint is to identify those factors that constrain or enhance organizational
effectiveness and to determine their implications. Environmental scanning is a systematic approach to detecting scientific, technical, economic, social, and political trends and events important to the organization. Porter’s (1980) Five Forces Model provides a general framework for performing the environmental scan in organizations. The five forces include the intensity of competitive rivalry, threat of substitutes, buyer power, supplier power, and barriers to entry. In addition to these, it is important to consider other environmental factors including sociocultural, political, legislative, and regulatory forces. Special attention should be given to the organization’s primary competitors, as this group is purposely acting to displace the organization’s position in the marketplace. Effective environmental scanning will enable the evaluator to anticipate changes emerging in the organization’s external environment. The consequences of this activity include fostering an understanding of the effects of external change on the organization and aiding in identifying contextual factors influencing organizational performance.

D. Conduct a SWOT analysis. The SWOT analysis focuses on both the organization’s internal and external environments. The external environment considered in the SWOT analysis is focused on specific elements that are identified as known opportunities and threats rather than scanning for broad environmental trends (outlined in the previous checkpoint). The purpose of a SWOT analysis is to systematically identify areas where the organization excels (i.e., strengths), areas of weakness, opportunities to leverage, and threats to mitigate. It is important to recognize that strengths and weaknesses are strictly internal to the organization. In other words, they are under the control of the organization. Opportunities and threats, in contrast to strengths and weaknesses, are external factors beyond the organization’s control. Some opportunities and threats may “fall out” of the environmental scan completed in the previous
checkpoint. However, attention should be given to searching for specific opportunities
and threats whose impacts on the organization are imminent.

Because this form of analysis is commonly used for strategic planning purposes
in organizational settings, an existing SWOT analysis may be available to the evaluator. If
this is the case, the current or past analyses should be carefully reviewed to acquire
greater insight into the essence of the organization. An existing SWOT analysis does not
relieve the evaluator of performing an updated analysis, but does facilitate the present
analysis by providing additional perspective. The results of the SWOT analysis facilitate
the identification of performance-level needs referred to in the next checkpoint.

E. Identify the performance-level needs of the organization. Each of the preceding
checkpoints provides the necessary information to identify the organization's
performance-level needs. Performance-level needs are needs that, if not met, lead to
dysfunction within the organization (Davidson, 2005). Identifying performance-level
needs reveals contextual evaluative criteria that can be added to the universal criteria of
merit for evaluating organizational effectiveness.

Step 3: Define the Criteria for the Evaluation

The OEC includes two categories of criteria: universal criteria of merit that apply
to all organizations and contextual criteria based on a performance needs assessment or
other relevant values (e.g., operative goals). The following checkpoints make up the heart
of this checklist. Criteria are necessary for credible evaluative conclusions to be made
about an organization's effectiveness. General guidelines for a list of criteria of merit
include seven specific requirements (Scriven, 1994, 2007a):
1. The list should refer to criteria and not mere indicators.

2. The list should be complete (i.e., no significant omissions).

3. The list should be concise.

4. The criteria should be nonoverlapping.

5. The criteria should be commensurable.

6. The criteria should be clear.

7. The criteria should be confirmable.

In addition to the general guidelines for developing criteria of merit noted above, specific guidelines for evaluating organization effectiveness can be appended to include these:

8. The criteria should be limited to the organizational level of analysis.

9. The criteria should reflect the relation between the organization and its environment.

10. The criteria should allow for the uniqueness of the organization.

11. The criteria should include both the means and ends of organizational activity.

12. The criteria should be stable yet provide the necessary latitude for organizational change and variability over time.

The starting point for any list of criteria of merit is an understanding of the nature of the evaluand and the properties that make it good, valuable, or significant (Scriven, 2005). In the case where the organization is the evaluand and the evaluation pertains to its effectiveness, the question to answer becomes, "What properties are included in the concept of organizational effectiveness?" As noted in the previous chapter, the operational definition of organizational effectiveness applied here is the
extent to which the organization provides sustainable value through the purposeful transformation of inputs and exchange of outputs, while minimizing harm from its actions.

Criteria, unlike indicators of merit, are definitionally connected to the evaluand (Scriven, 1959). They represent the properties or characteristics of a good or valuable evaluand. In contrast, indicators are empirically connected to a criterion variable (e.g., a correlate) and are susceptible to manipulation or can be easily corrupted (Scriven, 1991). Therefore, “the path of righteousness for evaluators is the path of criteria, not indicators” (Scriven, 2007a, p. 7).

The process used to identify the properties of an effective organization for inclusion in the OEC began with an understanding of the nature of the organization. Based on an organizational effectiveness literature review and nearly twenty years of professional business experience, a preliminary list of properties that define an effective organization was created. Specific attention was given to Scriven’s (2007a) guidelines for criteria of merit checklists. Most of the existing models for assessing organizational effectiveness do not include criteria of merit; those that do include criteria have generally limited the list to two or three criteria. However, Quinn & Rohrbaugh’s (1981) competing values framework (CVF) does offer such a list. The CVF offers a clear and commensurable list of criteria based on a factor analysis study using thirty criteria of merit that had been used in previous organizational effectiveness assessments. The elimination of overlapping criteria of merit and those that were not at the organizational level makes the CVF a well constructed criteria of merit list. Based on the rigorous efforts used to develop the CVF and the number of validation studies conducted (refer to Chapter 2 for additional details), the criteria of merit identified in the CVF provided a
foundation for the universal criteria of merit used in the OEC. However, the CVF was constrained by its dependence on the list of criteria selected for use in its factor analysis study, and it failed to identify the complete array of relevant value sources, including ethics. As a result, several omissions of relevant criteria were made, including:

1. No specific criterion for quality of output—the CVF assumes quality to be a facet of all criteria.

2. No specific criterion for fiscal health; that is, no cost associated with the organizational activity or financial considerations are included. The CVF assumes growth is a proxy for profitability.

3. No specific criterion for ethical or legal considerations.

The OEC addresses these omissions and includes a list of twelve criteria of merit for evaluating organizational effectiveness. The twelve criteria have been grouped into four dimensions to illustrate the connection with the operational definition of organizational effectiveness applied in this dissertation. The dimensions include (1) purposeful (efficiency, productivity, and stability); (2) adaptable (innovation, growth, and evaluative); (3) sustainable (fiscal health, output quality, information management, and conflict and cohesion); and (4) harm minimization (intra-organizational and extra-organizational). The twelve universal criteria of merit used in the OEC include:

1. **Efficiency.** A ratio that reflects the comparison of some aspect of unit performance to the costs (time, money, space) incurred for that performance. It is often used to measure aspects of a process other than just physical output.

2. **Productivity.** The ratio of physical output to input in an organization. Output consists of the goods or services that the organization provides, while the inputs include resources such as labor, equipment, land, facilities, etc. The more output produced relative to the input, the greater the productivity.
3. **Stability.** The maintenance of structure, function, and resources through time and more particularly, through periods of stress. This criterion is at the opposite end of the continuum of the innovation criterion (Campbell, 1977). Measures may include extent of job rotations, percent of processes routinized or formalized, and alignment of strategy and mission.

4. **Innovation.** The degree to which changes (either temporary or permanent) in process, procedures, or products are intentionally implemented in response to environmental changes (Price & Mueller, 1986).

5. **Growth.** The ability of an organization to import more resources than it consumes in order to maintain itself. Measures may include revenue and profit growth, change in workforce, net change in number of clients served, etc.

6. **Evaluative.** The extent to which an organization actively seeks out opportunities for improvement and incorporates feedback into its planning and operation processes to adapt to the internal and external environment. Measures may include percentage of new ideas attributed to internal sources; number of new initiatives launched; and use of benchmarking, best practices, or performance management systems.

7. **Fiscal health.** The financial viability of an organization as represented by its financial statements (e.g., balance sheet, income statement, cash flow statement). In most cases, ratios are used to allow for comparison with relevant organizations.

8. **Output quality.** The totality of features and characteristics of the primary good, service, or idea that bears on its ability to satisfy the stated or implied needs of the recipient. The output quality of the organization may take many operational forms, which are largely determined by the kind of good, service, or idea provided by the organization.

9. **Information management.** The completeness, efficiency, and accuracy in analysis and distribution of information. This includes cross-level collaboration, participative decision making, accessibility to influence, and communications inside and external to the organization (Campbell, 1977).

10. **Conflict and cohesion.** Cohesion is the extent to which organizational members work well together, communicate fully and openly, and coordinate their work
efforts. At the other end lies conflict, characterized by verbal and physical clashes, poor coordination, and ineffective dedication (Campbell, 1977; see also Price & Mueller, 1986).

11. **Intra-organizational harm minimization.** Those ethical and legal items that pertain to matters internal to the organization such as instances of ethical conduct breach, evidence of workforce training, and internal and external audits.

12. **Extra-organizational harm minimization.** Those ethical and legal items that pertain to matters external to the organization such as compliance with government regulations, ecological footprint, and internal and external audits.

The universal criteria of merit capture both the means and ends of organizational activities, allow for the comparative study of organizations or subunits, and recognize the uniqueness of the organization being evaluated. When combined with the contextual criteria of merit, the list provides an adequate inventory of the characteristics and properties of an effective organization.

In organizations where more than five participants will be involved in defining the criteria, measures, and performance standards, it is best to approach this step using a two-phase process. The first phase consists of using a small working group of two or three key informants to work through each of the checkpoints in Step 3. The evaluator and small working group complete the various checkpoints with the intent of developing a draft set of criteria, measures, and performance standards. This working group also identifies potential challenges or points of concern that may arise when the structured discussion is conducted with the larger group. The draft set of criteria, measures, and performance standards are provided to the primary beneficiaries or dominant coalition members in advance of the evaluator-facilitated workshop that occurs in the second phase of this step. Phase two consists of two 90-minute workshops intended to engage the primary beneficiaries or dominant coalition in an open, yet structured, dialogue and
exchange of ideas regarding organizational effectiveness and its dimensions. The checkpoints are used as the framework for the structured discussion and the draft set of criteria, measures, and importance weightings are adjusted as needed during this first review. The highly compressed time allotment for this workshop is intended to focus the discussion on the specific issues related to criteria development and requires the small working group to be well prepared prior to this workshop. It also requires that the evaluator "sell" the importance of the criteria of merit checklist in advance, because the immediate value of this exercise may not be seen by all participants. This is particularly true if an internal evaluator is used rather than an external evaluator. The second 90-minute workshop is used to review the performance standards and create performance matrices outlined in the last two checkpoints of Step 3.

A. Review universal criteria of merit of organizational effectiveness. The universal criteria of merit consist of the characteristics that define an effective organization. These characteristics are intended to be applicable to all organizations that are deliberately structured for a specific purpose. The universal criteria of merit are reviewed with the client to ensure each criterion and dimension is understood and to stimulate thinking about potential measures that may be used.

B. Add contextual criteria identified in the performance needs assessment. The information collected in the performance needs assessment may have revealed additional evaluative criteria that are unique to the organization. These criteria may result from the political, social, or cultural environment; developmental stage of the organization; current situational issues threatening the survival of the organization; or other matters that are unique to the organization at the particular time of inquiry. When considering contextual criteria of merit in multidivisional organizations, it is important to look across the
organization in addition to within the particular unit or division to ensure optimization in one unit does not result in suboptimization in another.

C. Determine the importance weightings for each criterion. The weighting of criteria by relative importance recognizes that some criteria are more important than other criteria. It also allows for a more complex inference (Scriven, 1994). Weighting is particularly important when the evaluative conclusions for each dimension must be synthesized into an overall evaluative conclusion regarding the organization's effectiveness. When using the OEC to conduct a formative evaluation that uses profiling to show how the organization is performing on the various effectiveness dimensions, weighting may be avoided; the client can use the multiple evaluative conclusions to identify and prioritize areas for improvement. However, when conducting a summative evaluation, it is necessary to go one step further than profiling and make an overall evaluative conclusion for the benefit of the client and the utility of the evaluation.

There are a number of different strategies for determining importance weightings, including voting by stakeholders or key informants, using expert judgment, and using evidence from a needs assessment, among others (Davidson, 2005). Two alternative methods for determining the relative important of criteria are offered here.

The first is a qualitative approach that consists of having the primary beneficiaries or dominant coalition agree on each criterion's importance using the categories of low, medium, high, and emergency. The emergency category is used to reflect the temporal importance or urgency of a criterion at a particular moment in time. It can be applied to criteria that may not be of high importance in the long term, but are of critical importance in the immediate term. For example, if an organization is suffering from cash flow or credit restrictions and only has the ability to pay its employees for
thirty days, growth and fiscal health criteria may be categorized as "emergency." Unless these areas are addressed with urgency, the organization's professional development programs or new product development rate become less relevant as the organization's survival is of immediate concern.

The second method is a quantitative approach using the analytic hierarchy process to derive a set of numeric weights from pair-wise comparisons (Saaty, 1980; see also Aczel & Saaty, 1983; Harker & Vargas, 1987). Each member of the dominant coalition identifies which of two performance criteria he or she believes is more important and then records the magnitude of the selected criterion's importance over the criterion not selected. This process is repeated until every criterion has been compared to the others. To determine the weightings, an organization-level matrix is created from each participant's pair-wise comparisons, the comparisons are normalized, and the importance weighting is calculated. This procedure can be done using spreadsheet software with intermediate-level knowledge of how to use the software.

Note that it is important to have all primary beneficiaries or dominant coalition members agree on the importance weightings in front of each other, before any data are collected. This "public" affirmation reduces the likelihood that participants will change their opinion after seeing the performance data or evaluative conclusions.

D. Identify observable and concrete measures for each criterion. The performance measures for each criterion are the factual data that will be collected and synthesized with the values (i.e., criteria) to produce the evaluative claims. It is best to use measures that can be observed and are stable and valid. Whenever possible, include several measures for each criterion, preferably from different sources. Of equal importance, agreement should be reached on precise nature of the measures and data sources. For
example, if revenues are a measure, is revenue recognized when the order is booked, invoiced, or money collected? Are revenues determined before or after the impact of sales incentive plans? Is the change in revenue determined year-over-year or quarter-over-quarter? In situations where performance cannot be observed or directly measured, the inclusion of multiple measures from different sources will increase the validity and credibility of the findings for the particular criterion and contribute differentially with unique information. When developing measures, it is important to focus on higher quality measures that are observable and credible rather than aiming for a high quantity of lower quality measures.

E. Identify performance standards for each criterion. Performance standards are the claims against which performance data are compared. In other words, standards are quality categories of increasing merit; in some organizations, they are referred to as benchmarks. The referents may include internal, peer-group, and best-demonstrated practices in any industry or segment, and in most cases organizations will have internal quality standards already in place that can be leveraged. However, it is appropriate to use industry or peer-group performance standards in addition to internal standards. In those cases where the peer group performs poorly, the use of best-demonstrated practices in any sector or industry as a referent is recommended. In the case of harm minimization, this dimension is an absolute measure and does not rely on comparisons to other organizations. In other words, one organization's illegal activities are not "less illegal" than another organization's illegal activities. Both have violated the law. A similar argument applies to ethical requirements.

For those criteria whose importance is identified as essential, a bar should be established to indicate the minimum level of acceptable performance. Below the bar, the
organization fails on that particular criterion. There are specific types of bars that can be used for this operation. A soft bar, for example, indicates a minimum level of acceptable performance for a particular dimension or subdimension of the evaluand to qualify for entry into a high-rating category. A global bar, sometimes referred to as a holistic bar, would be applied to those criteria where performance below a minimum level means the organization is ineffective overall, regardless of exemplary performance on the other dimensions (Davidson, 2005; Scriven, 1991). For instance, if a global bar has been established for specific ethical standards, an organization would be deemed ineffective if violations of ethical standards were discovered—no matter how well it performed on the other criteria of organizational effectiveness.

F. Create a performance matrix for each criterion. A performance matrix is a tool for converting descriptive data into an evaluative description or judgment. It can be used for determining absolute merit (i.e., grading) or relative merit (i.e., ranking). The most basic performance matrix includes a rating (e.g., excellent, fair, poor) and a description or definition of the rating. An example of a performance matrix to determine absolute merit for a specific criterion is shown in Table 2; a performance matrix to determine relative merit using a different criterion is shown in Table 3.
### Table 2

Performance Matrix for Determining Absolute Merit of the Management Support Criterion

<table>
<thead>
<tr>
<th>Rating</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Excellent</td>
<td>The organization requires only limited and infrequent support from its parent company or external consultants. It is capable of meeting its financial and operational goals with internal human resources and knowledge.</td>
</tr>
<tr>
<td>Good</td>
<td>The organization is self-sufficient, but requires some input from its parent company on specific issues on an infrequent basis.</td>
</tr>
<tr>
<td>Acceptable</td>
<td>The organization is self-sufficient, but requires some input from its parent company on specific issues on a regular (monthly) basis.</td>
</tr>
<tr>
<td>Marginal</td>
<td>The organization is self-sufficient but requires some input and assistance from its parent company on a frequent (several times monthly) basis. Financial and operational goals are somewhat of a challenge to meet without support from the parent company.</td>
</tr>
<tr>
<td>Poor</td>
<td>The organization lacks ability to self-manage without serious support from its parent company or external consultants. Few, if any, organizational goals can be achieved without support from its parent company.</td>
</tr>
</tbody>
</table>

### Table 3

Performance Matrix for Determining Relative Merit of the Profit Per Employee Criterion

<table>
<thead>
<tr>
<th>Rating</th>
<th>Net Profit Margin Percentile Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>Best</td>
<td>&gt; 95%</td>
</tr>
<tr>
<td>Better</td>
<td>75%-95%</td>
</tr>
<tr>
<td>Typical</td>
<td>50%-74%</td>
</tr>
<tr>
<td>Inferior</td>
<td>25%-49%</td>
</tr>
<tr>
<td>Worst</td>
<td>&lt; 25%</td>
</tr>
</tbody>
</table>

Creating performance matrices for each criterion forces organizational participants to think seriously about how they define performance, quality, or value (Davidson, 2005). Engaging primary beneficiaries or the dominant coalition in this process can result in increased buy-in for the evaluation, generate deeper interest in the evaluation process and outcomes, and increase the transparency of the evaluation. The definition of performance characteristics for each criterion can be accomplished during the second evaluator-facilitated workshop session in which each participant or group of
participants is assigned two or three criteria and works independently to develop performance matrices for them. Each participant or participant group then shares the proposed performance matrices with the entire group. Revisions are made, and a final version is accepted by the group. This cycle repeats until a performance matrix has been created for each of the criteria and subcriteria. It is advisable to fully develop all performance matrices prior to collecting any data to avoid the temptation to manipulate the evaluative descriptions and cut-offs to achieve positive ratings. This is particularly true for internal evaluators who may feel pressure from colleagues or superiors to produce a favorable evaluation.

Step 4: Plan and Implement the Evaluation

Items covered in this step focus on the evaluation plan and implementation. The primary determinants influencing the evaluation plan include evaluation team skills, organizational design and structure, available resources, and intended uses of the evaluation. The evaluation team (or individual) skills may narrow the evaluation design options, unless external resources are subcontracted. The organizational design, structure, and other contextual factors (e.g., culture and developmental stage) will also influence which type of data collection methods and sources are most appropriate. In some cases, participants will be willing and able to support data collection efforts. In most organizational evaluations, data sources are internal (e.g., organization members and archival records). The context will support certain data collection methods and inhibit others.

A. Identify data sources. Considering the measures identified in Step 3, the sources of data can be identified. Although many of the measures call for archival data stored as documents or records, the use of observations, surveys, and interviews with various
organizational participants will provide a fuller picture of the organization's performance and assist in uncovering side effects or side impacts that may not be revealed from archival data. If the organization being evaluated is a subunit in a vertically integrated organization, information from upstream units (i.e., those that provide components or service to the unit being evaluated) and downstream units (i.e., those that receive components or service from the unit being evaluated) will support triangulation of the data.

B. Identify data collection methods. The data collection methods are oftentimes directly influenced by a predetermined budget or a short timeframe for the evaluation. To address these influences, the data collection methods must be consistent with the purpose of the evaluation and needs of the organization, be flexible to take advantage of any data source that is feasible and cost-efficient, provide relevant information, and allow for comparisons from multiple data sources for purposes of triangulation (NPR, 1997; Yin, 2003). Every data collection method features some inherent form of measurement error, and using methods that have different types of bias guards against inaccurate conclusions. In addition, using multiple data collection methods and sources reduces the probability the results are due to artifacts of a given method, but represent a truer measure of organizational effectiveness. Hence, the credibility of the evaluation findings is strengthened when using multiple data collection methods and sources.

C. Collect and analyse data. Data collection activities should be nonintrusive and nondisruptive (to the greatest extent possible), cost-efficient, and feasible given the available resources. Although document and record review are likely to have the lowest impact on the organization, they should be managed to minimize frequency of requests and interruptions to the day-to-day activities of the organization's members. Awareness
of the how the evaluator may affect the organization should be considered whenever activities require the evaluator to interface with the organizational members or actors. For example, when collecting data from highly skilled technologists or professionals whose billable services can be charged out at $400 and up per hour, a 30-minute interview can carry quite a price—both in pecuniary terms and in opportunity costs. This illustration reinforces the point that care should be taken when developing the data collection methods and instruments.

Once the performance matrices have been defined and the data collected, data analysis begins. It is important to consider the client’s and other stakeholders’ understanding of and requirement for various types of data analysis. The primary consideration is that the results of the evaluation are clear, comprehensible, and credible to the client. For most organizational evaluations, qualitative and quantitative data will be collected. Regardless of the type of data collected, the analysis should be systematic and rigorous and will most likely include rich description of observed processes or interviews, as well as descriptive statistics that include measures of central tendency, variability, and relationships among variables.

*Step 5: Synthesize Performance Data with Values*

Synthesis is the combining of factual data and values into an evaluative conclusion. It is the final step in the logic of evaluation and is a primary distinction between evaluation and research. The synthesis process is what allows the evaluator to move beyond simply describing “what’s so” to answering the question, “So what?” (Davidson, 2005).
When the evaluation is used for formative rather than summative purposes, profiling the performance on the various dimensions of organizational effectiveness may be all that is required. The profile provides the client a graphical summary of performance and highlights areas to address to improve performance. For summative evaluations, synthesis of the various dimensions into an overall evaluative conclusion is required. In this case, the use of a performance matrix (described in Step 3) to synthesize the dimensions into an overall performance score or rating is recommended.

A. Create a performance profile for each criterion. The performance profile is a graphical illustration of how well the organization performs on each criterion according to the performance matrix. With the criteria listed on the vertical axis and the performance ratings shown on the horizontal axis, a performance bar extends outward to the appropriate rating. An example of a performance profile is shown in Figure 3. Note that the profiles for criteria are presented in a different color or shading from the dimension profile (shown as a solid bar). According to this profile, the organization is “good” on the workgroup cohesion and information management criteria, “acceptable” on the output quality criterion, and “marginal” on the fiscal health criterion. The rating (e.g., excellent, good, acceptable, marginal, or poor) for each criterion is determined by converting the criterion’s performance measures into a score and then calculating the average score for the criterion. For example, assume the rating excellent = 4.0, good = 3.0, acceptable = 2.0, marginal = 1.0, and poor = 0.0. If the performance on the two measures for the criterion was found to be good (3.0) and marginal (1.0) on another, the average criterion score is 2.33. No weighting is involved in this procedure. This step is repeated for each criterion until scores have been created for all criteria. To arrive at the
dimension profile, the importance weightings are applied to each criterion and the numerical weight-and-sum method is used to produce a dimension score.

When conducting a summative evaluation, the final synthesis step applies the same procedure used to arrive at the dimension profile, except in this case the importance weightings of dimensions are used and the numerical weight-and-sum procedure is used to arrive a composite score and overall grade of organizational effectiveness.

![Figure 3: Profile of Sustainable Dimension](image)

B. Create a profile of organizational effectiveness. In both formative and summative evaluations, profiling offers an easy-to-use and easy-to-understand tool for the client. When possible, including performance bars for both the dimensions and the criteria condenses broad insight in to a concise view. An example of a profile of organizational effectiveness is shown in Figure 4. In this profile, the organization is good to excellent on the purposeful and sustainable dimensions, but requires attention on the adaptable dimension. Defining and implementing actions that improve performance on the adaptable dimension would increase the organization's overall effectiveness.
C. Identify organizational strengths and weaknesses. Based on the organizational effectiveness profile, areas of strength and areas that require attention can be identified. This checkpoint is not intended to provide recommendations to the client. Rather, it is intended to highlight the organization's performance on the various dimensions. The appropriateness of the goal choices as well as performance against organizational goals identified in Step 1 may also be considered at this checkpoint. Inasmuch, the organizational effectiveness profile is used to support statements made regarding the degree of goal attainment. For example, if an organizational goal was to increase revenues by launching a new line of products in a new market and this goal was not achieved, the weak performance found on the innovation and growth criteria may be
used to illustrate why achievement of this goal was inhibited. Based on the findings and evaluative conclusions, the evaluator is in a unique position to report on the organization’s progress toward meeting its intended purpose or specific goals, and it is in this step that feedback regarding goal attainment is presented to the client.

In some cases, it may be useful to order the list of strengths and weaknesses so that the strongest or weakest dimensions are listed first. This allows the client to quickly grasp the areas of success and those requiring the most serious attention. When communicating organizational deficiencies, it is recommended to avoid personalization of the findings. This is not to suggest that negative findings should not be presented. Instead, it is to suggest that sensitivity must be shown and due consideration given to the level of evaluation anxiety that may be present. For example, using the phrase, “this is what the data tell us,” rather than “this is what we found” depersonalizes the findings and can facilitate constructive discussion.

Step 6: Communicate and Report Evaluation Activities

Communicating about the evaluation itself and reporting evaluation findings serve important roles within the organizational evaluation. During the evaluation, regular communications keep the client informed as to the status of the evaluation and can reduce negative reactions by the client and other organizational members. It can also provide reassurance that the costs (e.g., time, money, and other resources) being incurred are resulting in progress toward completing the evaluation. Reporting is generally thought of as the “end game” of the evaluation process. It is when evaluative conclusions are presented based on the synthesis of the factual data and values.
By and large, a long, written report will not be used in a business context. The client (or primary liaison) may read the entire report, but few others will. To ensure the findings are appropriately presented to others, the inclusion of a well written executive summary—one page preferred, two pages maximum—supported by a deck of slides using, for example, Microsoft’s PowerPoint® software is recommended. The slides should be numbered so it is evident to reviewers who are only seeing two or three slides, that more are available. From a communication style perspective, always summarize first, then explain. This is particularly important when communicating to executives. The executive summary should not be used as a “teaser”; rather it should give a summary of the evaluative conclusions and recommendations (if any). This is then followed by supporting evidence and explanations that led to the evaluative conclusions presented. The same approach should be followed in the written report; summary statements begin each section, followed by supporting evidence.

The highly charged nature of some evaluations, particularly those that involve summative decisions related to resource allocation or program continuance, require heightened awareness of the clients’ reactions. In some situations, the client may make inappropriate inferences from evaluative conclusions and recommendations. It is worthwhile to call attention to what the evaluative conclusions and recommendations imply—and even more importantly, what they do not imply. For example, if the evaluand was a pilot advertising program for a new line of furniture, and the program was determined to be cost-ineffective, a valid implication is that alternative programs are worth consideration. It does not imply that the entire marketing team is incompetent or that the product line should be abandoned. Clearly stating what is implied and what is
not implied facilitates a proper interpretation of the evaluative conclusions and may increase the utilization of the evaluation for the betterment of the organization.

A. Distribute regular communications about the evaluation progress. Communicating about evaluation activities on a regular basis serves two primary purposes: It keeps the need-to-know audience up-to-date on the evaluation’s progress and can help generate buy-in (Torres, 2001). The frequency and content of the communications should be based on the client’s preferences; a single approach may not suit all organizational evaluations or all audiences within a particular organization. In most cases, a one-page summary highlighting recent activities, status, upcoming activities, and potential obstacles can be issued biweekly to the client. The distribution may be via e-mail, uploaded to a secure Internet site for broader access, or both. Note that widely distributing information beyond the client and key stakeholders may be considered a political act and should be avoided. If an update is provided orally, it should be followed by a written summary for the benefit of both the client and evaluator.

B. Deliver a draft written report to client for review and comment. The draft written report is an important element in that it allows for clarifications, objections, and other comments to be received from the client prior to submitting a final report. By offering a draft for review, the evaluator is seeking evaluative feedback. This form of feedback performs two primary functions. First, it engages the client. This engagement encourages ownership of the report by the client and may increase the evaluation’s credibility and use. Second, findings that are not clearly communicated, are ambiguous, are erroneous, or need verification, may be discovered and corrected prior to the final written report. It is important to reflect carefully on the suggestions made by the reviewers. Care should
be taken to ensure that any changes included in the final report based on reviewer feedback do not distort the findings.

C. Edit the report to include points of clarification or reaction statements. Based on the client’s feedback to the draft report, the written report is edited as needed. In all cases, the opportunity for the client or a team member to include a reaction (i.e., rejoinder) statement in the final report should be offered. Although the inclusion of a reaction statement is more commonly found in politically charged or high-stakes environments, it is important to be sensitive to this issue to encourage inclusion of the voices of those who may be less powerful or who have valid arguments related to the evaluation findings. Note that the evaluator’s responsibility is not to solve political issues, but to ensure that viable points are raised.

D. Present written and oral reports to client. The final written report should be delivered to the client and a time scheduled to present the findings and evaluative conclusions. The focus of this checkpoint is on brevity and quality of content, rather than quantity. Moreover, the reporting format and delivery method should be chosen to maximize access to the findings (as appropriate), allow for client engagement, and be tailored to needs of various stakeholder groups. The use of quotes from interviews and tables, charts, and other visual tools help encourage audience interaction and support the verbal presentation (Torres, 2001). When communicating negative findings, it is important to stress the opportunity for organizational learning and improvement while minimizing “finger pointing” toward particular individuals or groups within the organization.

The inclusion of recommendations as part of the evaluation is generally expected in organizational settings. However, care should be taken to limit the recommendations
to operational recommendations that "fall out" of the evaluation and can be implemented with little or no extra cost to the client. These types of recommendations focus on the internal workings of the evaluand and are intended to facilitate improvement efforts. On the other hand, recommendations concerning the disposition of the evaluand (e.g., redirect resources from one business unit to another) are, in nearly all cases, inappropriate due to the evaluator's limited knowledge of the decision space (M. Scriven, personal communication, September 6, 2006). In those situations where the evaluator does have the required expertise and knowledge to make macro-recommendations, it should be made clear that a different type of evaluation is required—one that assesses alternative options for decision making (Scriven, 2007b).

E. Provide follow-on support as requested by client. Follow-on activities may include answering questions that arise after the client has taken time to absorb the findings, which may be some weeks or months following the delivery of the evaluation report. The evaluator should be available to answer questions, both technical and general in nature, to facilitate utilization of the evaluation (Coryn, 2006). Including a half-day follow-on session in the evaluation proposal and budget will allow the evaluator to further his or her utility to the client while not detracting from work on other funded projects. In addition, this follow-on support may offer the opportunity to incorporate ongoing performance monitoring and evaluation as a regular activity within the organization for continuous improvement and increased organizational effectiveness. From a project management perspective, it is important to ensure there is a sign-off from the client that the evaluation is complete so that follow-on work is explicitly covered by a different contract, in the case of an external evaluator, or within a new project scope, in the case of an internal evaluator.
Summary

The OEC presented in this chapter outlines a process framework to evaluate the unique expressions of organizational effectiveness. The distinction of this model from other models for evaluating organizational effectiveness lies in its explicit reliance on the definitional premise of organizational effectiveness based on the input-output cycle inherent in all organizations. Rather than basing the conceptualization of organizational effectiveness on the differences among various organization types, the OEC builds on the similarities of organizations. Another distinction is the utility for strategic and managerial decision making. The practical relevance of the OEC and its ease of use allow organizations to better understand which aspects of the organization need attention to ensure survival and viability of the organization.
CHAPTER IV

METHODOLOGY

Overview

Two separate investigations were conducted to empirically validate the OEC based on Stufflebeam's (2000) guidelines for developing evaluation checklists and Scriven's (2007a) paper on the logic and methodology of checklist. The first investigation was a nonexperimental study that sought input from subject matter experts (e.g., organizational and evaluation theorists) and targeted users (e.g., business consultants, managers, and professional evaluators) regarding the merit of the OEC. The study was intended to generate critical feedback on the OEC from the expert panel members and incorporate the feedback into a revised checklist. The second investigation applied the revised OEC to a real-world evaluation to assess the effectiveness of a for-profit organization. Both investigations were keyed to a common set of criteria to evaluate and improve the checklist.

Expert Panel Review

Participants

The expert panel was a purposive, convenience sample. It was a sample of convenience in that participants were selected based on availability and accessibility. It was purposive, however, in that the sampling procedure sought participants based on certain characteristics. More specifically, participants were selected based on a
demonstrated degree of expertise in terms of organizational research, evaluation, or senior-level management experience.

Organizational theorists recognized by the Academy of Management as *Distinguished Scholar Practitioner* award winners were identified as one group of potential candidates based on their successful application of theory or research in practice or their contribution to knowledge through extraction of learning from practice. In addition to the award winners, published scholars with research interests in the area of organizational effectiveness and evaluation theory were also considered potential expert panel members. Participants with evaluation expertise were recruited from the membership of the Business and Industry topical interest group of the American Evaluation Association (a professional society for evaluators). The Business and Industry interest group membership includes professional evaluators, consultants, and organizational researchers, among others, who have expertise in evaluation in business environments. Senior-level managers in for-profit and nonprofit organizations with at least ten years of management experience were identified using a combination of recommendations and personal contacts.

Nineteen individuals agreed to participate as expert panel members. Eight candidates declined the invitation to participate, and three did not reply to the invitation to participate in the study. Of the nineteen individuals who agreed to serve as an expert panel member, fifteen participated in the present study. Eight of the participants were female and seven were male. The majority of respondents (40%) were twenty-five to thirty-nine years of age. All participants indicated “a great deal” or “extensive” professional experience in research (60%), evaluation (73%), or management (73%). Figure 2 illustrates the participants’ extent of research, evaluation, and management
experience. Of the fifteen expert panel members, 80% held doctoral degrees, and 20% held masters degrees. The expert panel members represented a diverse array of professional and disciplinary areas of expertise, including organizational performance improvement, management, engineering, evaluation research, sociology, organizational development and training, education, change management, business consulting, strategic planning, leadership development, project management, marketing, and accounting.

Figure 5: Expert Panel Members’ Extent of Professional Experience in Research, Evaluation, and Management

<table>
<thead>
<tr>
<th></th>
<th>Research</th>
<th>Evaluation</th>
<th>Management</th>
</tr>
</thead>
<tbody>
<tr>
<td>None</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Very little</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Some</td>
<td>5</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>A great deal</td>
<td>7</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Extensive</td>
<td>2</td>
<td>6</td>
<td>6</td>
</tr>
</tbody>
</table>

Instrumentation

The purposes of the expert panel review were to provide critical feedback on the OEC as well as demonstrate content relevance and representativeness along with the substantive and consequential aspects of validity (Messick, 1975, 1980, 1995). To satisfy these purposes, three original instruments were used in the study: the OEC, a critical feedback survey, and sociodemographic questionnaire. The cover sheet attached to the
OEC included instructions to the expert panel member regarding the assessment procedure and completion of the surveys.

The OEC reviewed by the panel members consisted of three parts. Following the table of contents, the first part of the OEC included a one-page summary of the major steps and checkpoints. The second part of the checklist provided explanations of each checkpoint outlined in the summary page and included the criteria of merit for organizational effectiveness. The criteria of merit were grouped into four primary dimensions and eleven subdimensions. Suggested measures for each of the criteria were also included. The third part included a glossary of important terminology used throughout the OEC. The OEC document was formatted with extra-wide right margins to facilitate written feedback directly on the document.

The critical feedback survey (see Appendix B) was a pen-and-paper survey consisting of twenty closed-ended items and four open-ended questions. The survey was designed to collect specific data to evaluate the OEC based on Stufflebeam's (2000) eight criteria for evaluating a checklist including applicability to the full range of intended uses, clarity, comprehensiveness, concreteness, ease of use, fairness, parsimony, and pertinence to the content area. Each closed-ended item used an interval response format from 1 to 9, where 1 = strongly disagree and 9 = strongly agree. The rationale underlying this response format was that an expanded scale would allow respondents to more accurately express their positive or negative feelings about the OEC, thereby increasing the sensitivity to detect differences that may be muted when using a five- or seven-point scale or a dichotomous response format (e.g., 0 = disagree, 1 = agree). The four open-ended questions inquired about the key strengths and weaknesses of the checklist,
solicited perceptions of what was missing from the checklist, and requested suggestions or recommendations for improving the checklist.

The final instrument was a sociodemographic questionnaire (see Appendix C) used to gather information about the characteristics of the expert panel members, including gender; age group; highest educational degree attained; and the extent of professional research, evaluation, and management experience. An open-ended question provided the opportunity for respondents to clarify their specific discipline or area of expertise. The originally developed questionnaire consisted of six closed-ended items and one open-ended question. Like the Critical Feedback Survey, the sociodemographic questionnaire was administered to the panel members by means of a pen-and-paper survey.

Procedures

Based on the expert panel member selection criteria outlined earlier, recruitment letters were sent to potential expert panel members inviting them to participate in the study. The invitation letter described the study, explained what would be required of panel members, and invited the candidates to ask questions about participating in the study. Prior to participating in the study, participants were asked to read, sign, and return a consent form (see Appendix D) that provided a description of the study, its purpose, and an overview of the tasks they would be asked to complete. A unique four-digit code was assigned to each panel member and included on the OEC, Critical Feedback Survey, and sociodemographic questionnaire for tracking purposes.

Following their agreement to participate as expert panel members, the participants were sent the OEC, Critical Feedback Survey, and sociodemographic questionnaire via priority mail. All participants were asked to review the OEC for the
purpose of providing critical feedback, focusing on the strengths of the checklist and areas to improve, and identify items that were missing or confusing. Participants were encouraged to write directly on the checklist and were asked to complete the sociodemographic questionnaire and critical feedback survey after completing the checklist review. A postage-paid return envelope was included with the materials. Each participant was offered a copy of the revised OEC and a PowerPoint© presentation summarizing the findings.

All survey responses were coded (where necessary) as they were received, and entered into a database. As noted earlier, the critical feedback survey inquired about the OEC's applicability to the full range of intended uses, clarity, comprehensiveness, concreteness, ease of use, fairness, parsimony, and pertinence to the content area. The ratings for each item were analyzed using descriptive statistics. Responses to the open-ended questions in the critical feedback survey were placed in a table to facilitate content analysis that identified common criticisms, suggestions, and areas of strength.

Written comments were carefully reviewed using a hermeneutic interpretation process (Forster, 1994; Gadamer, 1989). The first step in this process was to extensively and intensively read and reread the written comments made on both the OEC and the Critical Feedback Survey. This was intended to move beyond a superficial understanding and identify underlying themes. Thematic analysis was used to cluster or impose meaning upon the themes—in effect, meaningfully reducing the data and identifying the respondents' perspective of the OEC (Lee, 1999; Neuendorf, 2002). This was followed by a comparison of the comments provided by each expert panel member, and the interpretations were triangulated across the multiple panel members to corroborate or falsify the imposed categories (Yin, 1994). In hermeneutical terms, the data were
recontextualized and reinterpreted (Krippendorff, 2004). Potentially conflicting changes to the checklist were seriously reflected upon, and the changes in combination were considered to the extent the checklist would be improved as a whole.

Following the analytical review, revisions were made to the checklist to ameliorate each identified problem. Consideration was given to potentially conflicting changes that could be made to the checklist to ensure they were mutually reinforcing and made sense for improving the checklist as a whole (Stufflebeam, 2000).

Field Study Using the OEC

Overview

An evaluation using the revised OEC to evaluate a for-profit organization’s effectiveness was conducted for the second investigation. The purpose of conducting a real-world evaluation was to answer the research question, “How practical is the OEC process for evaluating organizational effectiveness?”

Assessing practicality was done from both the client’s perspective and the evaluator’s perspective. The desire to incorporate contextual conditions (i.e., the organization’s environment) with a contemporary phenomenon (i.e., the OEC process) to answer the research question supported the use of an exploratory research design (Fink, 2003; Yin, 2003).

The proposition under consideration was that use of the OEC would facilitate the evaluation process and result in a valid, useable, credible, cost-effective, and ethical evaluation of organizational effectiveness. These primary criteria of merit for evaluations (Scriven, 2007b) are reflected in The Program Evaluation Standards (Joint Committee on

Evaluation Client

The criteria used to select the case study organization required that the organization be (1) a stand-alone entity or profit center (i.e., not a division or program within a larger organization), (2) in operation for at least three years, and (3) willing to provide access to internal company records and senior-level management. Several organizations were identified as potential candidates based on the selection criteria. Two organizations were contacted to gauge their level of interest in participating in the field study. After an affirmative response was received from one of the organizations, a meeting was scheduled with the organization’s senior management team to present an overview of the OEC and details regarding the field study purpose. Following the meeting, the organization agreed to participate in the study.

The participating organization was a privately held, for-profit company specializing in storytelling using a variety of print and electronic media. The U.S.-based company was founded in 2003 and operates a single location in the Midwest servicing business clients within a 100-mile radius of its office location. The organization is a limited liability company and has 16 full time employees, including the managing partners. As defined by the U.S. Small Business Administration (2008), the organization is a small business. The organization’s structure is moderately complex, yet highly formalized and centralized. The organization’s complexity is characterized by a modest degree of horizontal differentiation demonstrated by the organization’s functional specialization. Except for the highest management level, the span of control is narrow
with one or two subordinates reporting to a single manager or supervisor. However, the use of indirect reporting lines directly to the highest level of management suggests a flat organization structure with fluid communication channels. A high degree of formalization is found with the organization’s service offering (i.e., highly standardized outputs), yet the organization’s use of written rules and procedures is limited due to the highly skilled jobs that are performed by the employees. The organization’s high degree of centralization is expected based on its relatively low level of complexity and flat hierarchical structure.

The organization’s domain is characterized by its core service offering—writing and producing short stories. Its primary market segment is the funeral services market, while secondary market segments include assisted living centers and businesses. The organization’s ability to respond quickly with high quality stories and media production supported by a web-based infrastructure acts as a major differentiator from competitive offerings. Key aspects of the organizational domain include graphic design and media production expertise, story writing, and experiential offerings.

Procedures

The field study utilized the revised OEC to assess the client’s organizational effectiveness. The major steps included (1) establishing the boundaries of the evaluation, (2) conducting a performance needs assessment, (3) defining the criteria to be used, (4) collecting data, (5) synthesizing the performance data with the values, and (6) reporting the evaluation activities. The findings of the formative evaluation were intended to be used to identify areas for improvement and facilitate the prioritization of strategic
initiatives regarding organizational performance. Each step and checkpoint outlined in
the OEC was strictly followed to ensure fidelity to the OEC process.

Following the completion of the evaluation, a semi-structured interview was
conducted with the evaluation client to obtain the client’s perspective on the evaluation
with respect to its validity, credibility, utility, cost-effectiveness, and ethicality. In essence,
this was a formative metaevaluation intended to provide a modest level of assessment
that could be used to detect and correct deficiencies in the OEC process (Cooksy &
Caracelli, 2005). Adhering to the same rationale that Stufflebeam (2001) used when
assessing evaluation models, the OEC would be judged an unacceptable guide to
evaluation if it did not focus on the organization’s needs (utility), answer the questions it
was intended to address (validity and cost-effectiveness), present defensible conclusions
(credibility), and issue findings that were independent of the biases of the evaluation
participants (fairness and credibility).

The interview with the evaluation client was guided by a list of six open-ended
questions, including:

1. Did you find the evaluative conclusions to be believable? Why or why not? How
could credibility of the evaluative conclusions be improved?

2. Did you find the evaluative conclusions to be valid? Why or why not? How
could the validity be improved?

3. In terms of time, effort, and disruption to the organization, was the evaluation
cost-effective? Why or why not?

4. To what extent did you find the evaluation to be fair and complete, ethical, and
show due regard for those involved in the evaluation?

5. Overall, to what extent did you find the evaluation to be useful? Why or why
not? How could the utility be improved?

6. What actions, if any, do you expect the organization will take because of this
evaluation?
Because the study’s focus was on assessing the practicality of the OEC in hopes of improving the tool, a conscious decision was made by the investigator to focus the metaevaluation specifically on the evaluation findings and process from the client’s perspective in an effort to identify strengths and weaknesses of the OEC process. Issues and opportunities that were discovered by the evaluator during the field study were systematically noted and addressed by revising or clarifying the OEC where appropriate.

Aspects of Reliability and Validity

It was necessary to establish two types of reliability in the first study: rating and internal consistency. The use of a systematic approach for the expert panel members to apply the ratings minimized the measurement error. This established rating reliability. Cronbach’s coefficient alpha ranged from .79 to .96 for each construct in the critical feedback survey. This suggests a high estimate of reliability based on the mean inter-item correlation across the panel members. The use of multiple expert panel members with diverse backgrounds further implies internal consistency and potential stability over time (Yin, 2003). With respect to the content analysis performed on the qualitative feedback, the unit of analysis was a character string that referred to a specific idea (e.g., a word). This is “the smallest and, as far as reliability is concerned, the safest recording unit for written documents” (Krippendorff, 2004, p. 104).

Validity is an approximation of the truth of a given proposition, inference, or conclusion. More specifically, it is an evaluation argument about the extent to which theory and empirical evidence support the appropriateness of interpretations and inferences based on scores or other modes of assessment (Cronbach, 1988; Messick,
1989). With respect to the OEC, the aspects of construct validity (Messick, 1995) addressed included content, substantive, generalizability, external, and consequential.

The content aspect of construct validity includes evidence of content relevance and representativeness. Both of these elements were appraised by professional judgments made by the expert panel members using the critical feedback survey and captured in the written comments made directly on the OEC document. Substantive validity, the extent to which items in the OEC can be accounted for specifically in the context of the environment, was addressed by the expert panel review and through the field study where the contextual element was prominent (Loevinger, 1957; Messick 1995).

The generalizability aspect of construct validity in this study focused on analytic generalizability rather than theoretical or statistical generalizability (Kvale, 1996; Yin, 2003). It more traditional terms, analytic generalizability is akin to ecological validity. A single-case field-study design was used to examine the application of the OEC in a for-profit organization in an attempt to generalize a particular set of results to the broader theory that the OEC can be applied to any organization. The generalizability evidence is based on the fundamental argument that all organizations are social units that perform the same basic function. That is, inputs are transformed into outputs and exchanged with the environment with the intent of attaining a specific goal or fulfilling a purpose. Regardless of the type of organization (e.g., for-profit or nonprofit), its structure (e.g., centralized or decentralized), or its purpose, all organizations attempt to survive and maximize returns using the input-output cycle described in Chapter 1. Because of this common activity, the use of a single-case study supported by expert panel feedback
pertaining to applicability of the OEC is used as evidence for the (analytic) generalizability aspect of construct validity.

The external aspect of construct validity was considered in terms of the extent to which the OEC was consistent with other empirically validated models to assess organizational effectiveness. Of the influential models developed, only the CVF (competing values framework) specifically included criteria of merit for evaluating organization effectiveness (Quinn & Rohrbaugh, 1981). The CVF is also one of the few models whose psychometric properties have been examined using multitrait-multimethod analysis and multidimensional scaling (Quinn & Spreitzer, 1991) and structured equation modeling (Kalliath, Bluedorn, & Gillespie, 1999). As outlined in Chapter 3, the OEC includes all of the criteria used in the CVF in addition to several other criteria added to address the ethical and quality dimensions not specifically captured in the CVF.

The final facet of construct validity addressed in this investigation was the consequential aspect. This aspect includes evidence and rationale for evaluating the intended and unintended consequences of interpretation and use (Messick, 1995). Because the OEC is intended to benefit both users (e.g., professional evaluators, organizational consultants, managers) and organizations that are evaluated, it is important to accrue evidence of such positive consequences. The expert panel members provided feedback regarding the utility of the OEC in the Critical Feedback Survey. This evidence provides a measure of value from the perspective of the tool’s potential users. The field study also provided evidence of the OEC’s utility (i.e., practicality and value) from an organization’s perspective. Both of these investigations provide evidence and rationale for the consequential aspect of construct validity as defined earlier.
CHAPTER V

EXPERT PANEL AND CASE STUDY FINDINGS

Expert Panel Critical Review

A total of fifteen expert panel members participated in the critical review of the OEC. However, one of the respondents provided only written comments and did not complete the critical feedback survey. All of the closed-ended survey findings presented herein are based on fourteen respondents. The fourteen respondents included six senior executives and organizational consultants (referred to collectively as managers in this chapter), four professional evaluators, and four organizational or evaluation theorists. Descriptive data summarizing the ratings for each criterion used for evaluating the checklist are presented in Table 4. Across all panel members, the highest rating was found with the fairness criterion ($M = 7.05, SD = 1.52$). This was followed by concreteness ($M = 6.96, SD = 1.69$) and clarity ($M = 6.93, SD = 1.59$). The lowest-rated criterion was ease of use ($M = 5.61, SD = 1.68$). The primary role of each respondent (i.e., professional evaluator, manager, or theorist) reveals distinct perspectives held by each group with respect to the OEC and its performance against the criteria. Among the three groups, evaluators found the checklist to perform well on clarity ($M = 7.92, SD = .50$) and concreteness ($M = 7.13, SD = 1.18$) criteria, while parsimony ($M = 5.42, SD = 2.50$) received the lowest mean rating among this group. Note that parsimony captures elements specific to the cost-effectiveness of the OEC.
Table 4
Critical Feedback Survey Ratings by Respondent Role

<table>
<thead>
<tr>
<th>Construct</th>
<th>Overall</th>
<th>Evaluator</th>
<th>Manager</th>
<th>Theorist</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$M$</td>
<td>$SD$</td>
<td>$Min$</td>
<td>$Max$</td>
</tr>
<tr>
<td>Fairness</td>
<td>7.05</td>
<td>1.52</td>
<td>3.67</td>
<td>9.00</td>
</tr>
<tr>
<td>Concreteness</td>
<td>6.96</td>
<td>1.69</td>
<td>2.50</td>
<td>9.00</td>
</tr>
<tr>
<td>Clarity</td>
<td>6.93</td>
<td>1.59</td>
<td>3.33</td>
<td>9.00</td>
</tr>
<tr>
<td>Pertinence to the content area</td>
<td>6.86</td>
<td>1.69</td>
<td>2.00</td>
<td>9.00</td>
</tr>
<tr>
<td>Applicability to full range of intended uses</td>
<td>6.52</td>
<td>1.79</td>
<td>2.00</td>
<td>9.00</td>
</tr>
<tr>
<td>Comprehensiveness</td>
<td>6.46</td>
<td>1.80</td>
<td>2.50</td>
<td>8.50</td>
</tr>
<tr>
<td>Parsimony</td>
<td>5.70</td>
<td>1.71</td>
<td>2.00</td>
<td>8.00</td>
</tr>
<tr>
<td>Ease of Use</td>
<td>5.61</td>
<td>1.68</td>
<td>2.50</td>
<td>8.00</td>
</tr>
<tr>
<td>Total</td>
<td>6.51</td>
<td>1.36</td>
<td>3.58</td>
<td>8.04</td>
</tr>
</tbody>
</table>

Note. Response format from 1 to 9, where 1 = strongly disagree and 9 = strongly agree. Internal consistency for each construct was measured using Cronbach's alpha. The critical feedback survey instrument is located in Appendix B.
Managers also rated the OEC highly on the concreteness criterion \( (M = 7.83, SD = .82) \) and found the pertinence to the content area criterion to be of similar merit \( (M = 7.83, SD = .68) \). The ease of use criterion was rated the lowest by managers \( (M = 5.58, SD = 1.93) \). The third group of respondents— theorists—gave the fairness criterion the highest ratings \( (M = 6.83, SD = 2.15) \) and parsimony and ease of use the lowest ratings of the eight criteria. Table 5 shows each of the criteria in rank order for each group.

<table>
<thead>
<tr>
<th>Evaluators</th>
<th>Managers</th>
<th>Theorists</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Clarity</td>
<td>1. Concreteness</td>
<td>1. Fairness</td>
</tr>
<tr>
<td>2. Concreteness</td>
<td>2. Pertinence to content area</td>
<td>2. Applicability to full range of intended uses</td>
</tr>
<tr>
<td>3. Pertinence to content area</td>
<td>3. Comprehensiveness</td>
<td>3. Comprehensiveness</td>
</tr>
<tr>
<td>5. Fairness</td>
<td>5. Clarity</td>
<td>5. Pertinence to content area</td>
</tr>
<tr>
<td>6. Ease of use</td>
<td>6. Applicability to full range of intended uses</td>
<td>6. Clarity</td>
</tr>
</tbody>
</table>

The Shapiro-Wilk test was used to test the null hypothesis that the data for each group were normally distributed. At the \( p < .05 \) significance level, the null hypothesis was not rejected (Evaluator: \( W = .91, p = .49 \); Manager: \( W = .86, p = .18 \); Theorist: \( W = .93, p = .57 \)), and reasonable normality was assumed for further analysis. A one-way
analysis of variance was conducted to explore the difference in the critical feedback survey mean total scores among evaluators, managers, and theorists. The test for homogeneity of variances was satisfied using Levene’s test ($p = .35$), and the assumption of equal variances was not violated. As shown in Table 6, no statistically significant difference was found among the group means [$F(2, 11) = 1.54, p = .26$]. However, the power to detect a difference was small at .33 (Faul, Erdfelder, Lang, & Buchner, 2007). The effect size, calculated using eta squared, was large ($\eta^2 = .22$) as interpreted using Cohen’s (1988) guidelines for effect size.

Table 6

<table>
<thead>
<tr>
<th></th>
<th>DF</th>
<th>$SS$</th>
<th>$MS$</th>
<th>$F$</th>
<th>$p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between Groups</td>
<td>2</td>
<td>5.27</td>
<td>2.63</td>
<td>1.54</td>
<td>0.26</td>
</tr>
<tr>
<td>Within groups</td>
<td>11</td>
<td>18.79</td>
<td>1.71</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>13</td>
<td>24.05</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Due to the small sample size, the normality test may have failed to detect nonnormality of the distribution. Assuming this is the case and the data are not normally distributed, the Kruskal-Wallis H test can be used as the nonparametric alternative to the one-way between-groups analysis of variance (Hollander & Wolfe, 1999). However, the Kruskal-Wallis H test also indicated no statistically significant difference across the three groups [$\chi^2(2) = 2.44, p = .30$].

A one-way between-groups analysis of variance was performed to explore the impact of role (i.e., evaluator, manager, and theorist) on the score for each construct, as measured by the total mean score for the particular criterion. The assumption of equal variances was violated in two cases—clarity and comprehensiveness. In these cases, the
Welch test was used to compare the equality of means. Table 7 presents the results of the one-way analysis of variance and includes the effect size, calculated as eta squared, for each construct. Table 8 presents the robust test for equality of means using the Welch test. No statistically significant difference was found among the group means for any construct measured at the $p < .05$ significance level. Effect sizes ranged from $\eta^2 = .02$ to $\eta^2 = .39$. When converted to Cohen's $f$ value, the effect sizes ranged from .15 to .80 (Cohen, 1988, p. 283).

### Table 7

<table>
<thead>
<tr>
<th></th>
<th>$F$</th>
<th>$p$</th>
<th>$\eta^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Applicability</td>
<td>.10</td>
<td>.91</td>
<td>.02</td>
</tr>
<tr>
<td>Concreteness</td>
<td>3.03</td>
<td>.09</td>
<td>.35</td>
</tr>
<tr>
<td>Ease of Use</td>
<td>.52</td>
<td>.61</td>
<td>.09</td>
</tr>
<tr>
<td>Fairness</td>
<td>.45</td>
<td>.65</td>
<td>.08</td>
</tr>
<tr>
<td>Parsimony</td>
<td>.78</td>
<td>.48</td>
<td>.12</td>
</tr>
<tr>
<td>Pertinence</td>
<td>3.51</td>
<td>.07</td>
<td>.39</td>
</tr>
</tbody>
</table>

### Table 8

<table>
<thead>
<tr>
<th></th>
<th>$F'$</th>
<th>$p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clarity</td>
<td>3.71*</td>
<td>.09</td>
</tr>
<tr>
<td>Comprehensiveness</td>
<td>2.90*</td>
<td>.16</td>
</tr>
</tbody>
</table>

* Asymptotically $F$ distributed
Qualitative Feedback: Narrative Critique

In addition to the closed-item questions, the critical feedback survey asked the expert panel members four open-ended questions:

1. What are the key strengths of the checklist?
2. What are the key weaknesses of the checklist?
3. What is missing from the checklist?
4. What would you suggest or recommend improving?

The findings from the content analysis of the narrative critiques are presented below. The recording units of analysis included any character string that referred to a specific idea based on categorical distinctions. In all cases, the recording unit consisted of a single word (e.g., comprehensive) or phrase (e.g., clearly outlined process). The entire written response for the particular question made up the context unit. The small sample size (N = 15) of expert panel members providing qualitative feedback allowed for the analysis of all written comments made on the survey form. In other words, the entire universe of text of was analyzed.

A total of 139 recording units were classified into one of eight categories. Of these 139 units, 71 were listed as key strengths of the checklist, 29 as key weaknesses, 20 as items missing from the checklist, and 19 as items to improve. The categories used for classification purposes included the same eight criteria used in the quantitative analysis: applicability to the full range of intended uses, clarity, comprehensiveness, concreteness, fairness, ease of use, parsimony, and pertinence to the content area. A representative sample of recording units and their classifications are shown in Table 9.
Table 9

Representative Sample of Recording Units and Classifications

<table>
<thead>
<tr>
<th>Recording unit</th>
<th>Classification</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Strength</strong></td>
<td></td>
</tr>
<tr>
<td>Clearly defined steps to a rational end.</td>
<td>Clarity</td>
</tr>
<tr>
<td>Brevity.</td>
<td>Parsimony</td>
</tr>
<tr>
<td>Sound theory.</td>
<td>Concrete</td>
</tr>
<tr>
<td>Can be applied by an internal and/or external evaluator.</td>
<td>Applicability</td>
</tr>
<tr>
<td>Flexibility.</td>
<td>Ease of use</td>
</tr>
<tr>
<td><strong>Weakness</strong></td>
<td></td>
</tr>
<tr>
<td>Based on the audience, I believe more details/examples may have been useful to illustrate particular concepts.</td>
<td>Clarity</td>
</tr>
<tr>
<td>Response cost for client may be high.</td>
<td>Parsimony</td>
</tr>
<tr>
<td>Checklist requires modifications for use in governmental and other not-for-profit organizations.</td>
<td>Applicability</td>
</tr>
<tr>
<td>Would require an evaluator with good people skills to get the organization to provide the data plus benchmarks.</td>
<td>Ease of use</td>
</tr>
<tr>
<td>Too little information in places (p 14).</td>
<td>Comprehensiveness</td>
</tr>
<tr>
<td><strong>Missing</strong></td>
<td></td>
</tr>
<tr>
<td>Diversity of approaches and perspectives.</td>
<td>Pertinence</td>
</tr>
<tr>
<td>Context valuables.</td>
<td>Applicability</td>
</tr>
<tr>
<td>More about the concept of criteria of merit would be desirable.</td>
<td>Clarity</td>
</tr>
<tr>
<td>Issues of power and politics.</td>
<td>Fairness</td>
</tr>
<tr>
<td>Any time a glossary term appears in the checklist, you should put in all caps and bold it for the users to easier understand that there is a specific definition.</td>
<td>Ease of use</td>
</tr>
</tbody>
</table>
Table 9—Continued

<table>
<thead>
<tr>
<th>Recording unit</th>
<th>Classification</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Improve</strong></td>
<td></td>
</tr>
<tr>
<td>Beef-up and provide more thorough description of the criteria of merit.</td>
<td>Comprehensiveness</td>
</tr>
<tr>
<td>More illustrations of concepts using examples.</td>
<td>Clarity</td>
</tr>
<tr>
<td>While it includes some measures that might be cross-sector, that would still require some more work for evaluating nonprofits.</td>
<td>Applicability</td>
</tr>
<tr>
<td>A paragraph on the importance of stakeholder input into the determination of values use in the evaluation.</td>
<td>Clarity</td>
</tr>
<tr>
<td>The overview provided a contextual foundation, but no insight into the direction, focus, or summary of the OEC. Perhaps something like a one-page executive summary that highlights/briefly explains the [steps] presented.</td>
<td>Clarity</td>
</tr>
</tbody>
</table>

The number of strengths included in the narrative critique section of the survey represented 51 percent of the total recording units. As shown in Figure 6, the largest number of strengths identified by the expert panel members concerned the clarity of the OEC (18%). This was followed by comprehensiveness (16%), ease of use (16%), concreteness (14%), fairness (10%), pertinence to the content area (10%), applicability to the full range of its intended uses (8%), and parsimony (8%).

The number of weaknesses included in the narrative critique section of the survey represented 20 percent of the total recording units. Figure 7 enumerates the weaknesses in each category identified by the expert panel members. The largest number of weaknesses identified by the expert panel members concerned the clarity of the OEC (25%). This was followed by ease of use (17%), parsimony (17%), comprehensiveness (10%), fairness (10%), applicability (7%), concreteness (7%), and pertinence (7%).
Figure 6: Number of Strengths Identified by Category and Primary Role

Figure 7: Number of Weaknesses Identified by Category and Primary Role
The number of recording units captured in the responses to the question, "What is missing from the checklist?" represented 14 percent of the total units. As shown in Figure 8, the largest number of missing items identified by the expert panel members concerned the clarity the OEC (45%). This was followed by applicability to the full range of its intended uses (25%) and ease of use (15%).

The number of suggestions or recommendations for improvement included in the narrative critiques represented 14 percent of the total recording units. Figure 9 illustrates the number of improvements in each category suggested by the expert panel members. The largest number of suggestions concerned the clarity of the OEC (42%). This was followed by applicability to the full range of intended uses (26%) and ease of use (11%).

![Figure 8: Number of Items Missing Identified by Category and Primary Role](image-url)
Observations from Survey Responses

In an effort to compare the responses from the closed-ended items and open-ended narrative feedback, a net strength score was created for the qualitative feedback. Net strength is the difference between the normalized frequencies for each criterion’s identified strengths and weaknesses. The net strength score provided a ranking of the perceived strengths of the OEC based on each criterion. Table 10 presents the frequency of identified strengths and weaknesses for each construct, the normalized values, and the net strength score. The rank order of each construct based on the net strength score can be compared to the rank order of each construct based on the mean scores from the closed-ended survey items.
<table>
<thead>
<tr>
<th></th>
<th><strong>Strength</strong></th>
<th></th>
<th><strong>Weakness</strong></th>
<th></th>
<th><strong>Net Strength</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Frequency</td>
<td>Normalised</td>
<td>Frequency</td>
<td>Normalised</td>
<td></td>
</tr>
<tr>
<td>Applicability</td>
<td>6</td>
<td>8</td>
<td>2</td>
<td>7</td>
<td>1</td>
</tr>
<tr>
<td>Clarity</td>
<td>13</td>
<td>18</td>
<td>7</td>
<td>24</td>
<td>-6</td>
</tr>
<tr>
<td>Comprehensiveness</td>
<td>11</td>
<td>15</td>
<td>3</td>
<td>10</td>
<td>5</td>
</tr>
<tr>
<td>Conciseness</td>
<td>10</td>
<td>14</td>
<td>2</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>Ease of use</td>
<td>11</td>
<td>15</td>
<td>5</td>
<td>17</td>
<td>-2</td>
</tr>
<tr>
<td>Fairness</td>
<td>7</td>
<td>10</td>
<td>3</td>
<td>10</td>
<td>0</td>
</tr>
<tr>
<td>Parsimony</td>
<td>6</td>
<td>8</td>
<td>5</td>
<td>17</td>
<td>-9</td>
</tr>
<tr>
<td>Pertinence</td>
<td>7</td>
<td>10</td>
<td>2</td>
<td>7</td>
<td>3</td>
</tr>
</tbody>
</table>

The net strength score is plotted against the mean score to produce the output shown in Figure 10. Concreteness and pertinence were ranked in the top half of both the qualitative and quantitative responses, while ease of use and parsimony were ranked in the bottom half of both sets of responses. Although high mean scores for fairness and clarity placed these two criteria in the top half of the quantitative responses, they were in the lower half of the qualitative responses based on the net strength score. With respect to comprehensiveness and applicability to the full range of intended uses, both were ranked in the bottom half of the quantitative survey mean scores. However, both of these criteria were ranked in the top half of the qualitative responses based on the net strength score. This integrated perspective suggests that the OEC performed best on concreteness and pertinence to the content area, and was weakest on ease of use and parsimony.
Qualitative Feedback: Written Comments

In addition to completing the critical feedback survey, expert panel members were asked to write comments directly on the OEC document. A total of ninety-seven remarks were handwritten on the checklists; an additional thirty-one were typewritten and submitted with the checklist document and surveys.

The unit of analysis for these written comments was the complete sentence or group of sentences making up the entry. The comments were analyzed and classified into one of six categories: (1) minor content, (2) major content, (3) formatting, (4) grammar, (5) structural, and (6) other. A minor content classification was assigned when the written comment suggested a clarification of terminology or phrasing that was limited to one or two sentences. When a suggested edit required a new paragraph or the editing of an existing paragraph or multiple paragraphs, it was classified as major content. Notations reflecting formatting (e.g., italics, bold face type, emphasis) were classified as
formatting whereas punctuation, syntax, and semantic edits were classified as grammar. The fifth category used in the classification schema was structural. Suggestions that required reorganization of the checklist—such as creating a new checkpoint, reordering checkpoints, or moving checkpoints to a different location within the checklist—were classified as structural. The final category, other, was used to classify all other entries not specifically captured by one of the other five categories. An example of a verbatim comment associated with each type of classification is shown in Table 11.

The majority of written comments were classified as minor content (52%). This was followed by the categories other (24%), structural (9%), formatting (5%), major content (5%), and grammar (5%) as shown in Figure 11.

Table 11
Sample Verbatim Written Comments and Classifications

<table>
<thead>
<tr>
<th>Written comment</th>
<th>Classification</th>
</tr>
</thead>
<tbody>
<tr>
<td>It might be helpful to discuss that when internal persons are used, the evaluator may have to factor in a longer timeframe since these individuals may still have to perform their substantive duties along with the new assignment. Also, they may require special training.</td>
<td>Minor content</td>
</tr>
<tr>
<td>The last few paragraphs of section five appear to be the same as what was stated in section three.</td>
<td>Major content</td>
</tr>
<tr>
<td>The various subgroups in this section are making me wish you had italicized all of these new terms such as dominant coalition.</td>
<td>Formatting</td>
</tr>
<tr>
<td>Stable is not synonymous with reliable.</td>
<td>Grammar</td>
</tr>
<tr>
<td>The creation of performance matrices should have been done before data collection. I would put it at the last step, under step three.</td>
<td>Structural</td>
</tr>
<tr>
<td>This was very well done. A thorough description of the OEC.</td>
<td>Other</td>
</tr>
</tbody>
</table>

129
Discussion of Expert Panel Findings

Although the analysis did not identify a statistically significant difference among the total mean scores, the practical and meaningful significance of the quantitative data was enhanced by the qualitative feedback collected from both the critical feedback survey and the written responses made on the checklist. Inasmuch, the purpose of the analysis was not to test any predictions or hypotheses, but to aid with the development and refining of the OEC. Effect sizes were found to be large with several criteria, but the power to detect any significant difference was low. Stevens (1996) suggests that when small group sizes (e.g., N = 15) are involved, one should consider increasing the significance level to compensate for insufficient power. However, the slight increase in power from .33 to .47 resulting from a change in the significance level from .05 to .10 did not provide value for this particular study. Furthermore, the increased possibility of a
Type I error, supplemental qualitative data, and emphasis on practical significance provided additional support for the decision not to increase the significance level. As a result, the interpretation of the statistically nonsignificant findings is treated with care.

A number of issues surfaced with respect to the strengths and weaknesses of the OEC based on the quantitative and qualitative feedback received from the expert panel. In general, the OEC performed well with respect to its pertinence to organizational effectiveness evaluation, clarity, fairness, and sound theory. Specifically, the high scores on fairness, concreteness, pertinence to the content area, and applicability to the full range of intended uses provide evidence for the content relevance and representativeness, as well as substantive validity, of the OEC. Although the feedback from senior executives and business consultants serving on the expert panel was critical, it was favorable. The relatively high scores on the critical feedback survey and the written comments made on the checklist suggest this group found the OEC to be relevant, sensible, and comprehensive. For example, one senior executive commented, “Clearly outlined process—thorough process with reasonable guidance for such short paragraphs.” Another stated, “It exists as a complete, comprehensive package in a relatively compact form.”

The critical feedback survey responses from professional evaluators were found to be on par with the responses from senior executives and business consultants. This finding is of practical significance, in that the OEC is designed for and intended to be used by professional evaluators, managers, and business consultants. The organizational and evaluation theorists who reviewed the OEC assessed the checklist favorably with regard to fairness and applicability to organizational evaluation. However, in contrast to the professional evaluators and managers serving on the expert panel, theorists found
the OEC's clarity and relevancy to be relatively weak. For example, one theorist stated, "It is quite techno-rational [and] implies a rather positivist, reductionist approach that I am not convinced of in today's rapidly changing and knowledge environment." Perhaps of greater consequence with respect to improving the OEC, all three groups were in agreement on the need to address the checklist's ease of use and parsimony.

Although no statistically significant difference was found among the critical feedback survey mean scores of the three groups (i.e., evaluators, managers, and theorists), observations of the mean scores suggest the tool was well received by evaluators and managers and, to a lesser degree, by theorists. The qualitative data collected from the hand-written comments by the expert panel members tend to support this inference. For example, one professional evaluator noted, "The checklist represents a sound evaluation flow that integrates the interests of organizational effectiveness as outlined in the article. Well done." In another case, an organizational consultant commented, "I like this [criteria] checklist—it is a good tool for finding key performance indicators." Likewise, a theorist commented, "Good systematic manner to address evaluation of OE [organizational effectiveness]." These statements, in addition to other qualitative data collected, suggest positive consequences of the OEC and its use. What's more, the diversity of the panel (described in Chapter 4) supports the general notion of reasonable applicability and acceptance of the OEC across potential users having different professional interests and backgrounds.

Revisions to the Original OEC

Based on the findings outlined earlier, a number of changes were made to the OEC in an effort to build on the strengths of the checklist and address the deficiencies
identified by the expert panel members. As a practical matter, all suggestions regarding edits to grammar (e.g., punctuation, syntax, and semantics) were addressed as appropriate. A keen focus was placed on identifying opportunities to modify the OEC to improve its ease of use and parsimony. The edits made to address these two aspects consisted primarily of incorporating clarifying terms and explicatory descriptions to each checkpoint. In addition, the number of examples included throughout the checklist was increased to elucidate particular checkpoints and the overall process. The additional examples and clarifications increased the checklist’s word count by approximately 25 percent or 2,250 words.

The most significant structural change made to the OEC was specifically intended to facilitate the efficient use of the checklist and increase its parsimony. As noted earlier, parsimony was purposely intended to capture the cost-effectiveness of the checklist. Costs in this scenario include time, resources (both pecuniary and nonpecuniary), specialist expertise, and the cost of the evaluator. The structural change amalgamated the distinct components of Step 5 into Step 3 of the process and resulted in reducing the number of steps in the OEC from seven to six. It also eliminated overlapping activities found in Step 3 and Step 5 that contributed to the perception of the checklist as excessively complex, as evidenced by the statements, “Requires some considerable sophistication to fully comprehend it” and, “It does not seem easy to apply.” Another structural change was made in Step 1 by expanding the number of checkpoints from six to seven. The additional checkpoint, conduct an evaluability assessment, was originally a subcomponent of another checkpoint. The addition of this checkpoint provides for increased clarity and aids comprehension. Other changes made to Step 1 were minor content edits, added examples, and a reordering of the checkpoints.
to improve efficiency of use. More specifically, the additional content provided an expanded overview of organizational domain, clarification of organizational participants, and a notation regarding the consequences of using organization members in a collaborative effort.

Step 2 was modified to include a separate checkpoint for assessing the internal knowledge needs of the organization, and three of the five checkpoints were rewritten to improve clarity. A recommended change to Step 2 that was not implemented was the suggestion to eliminate the specific reference to a SWOT analysis and leave the approach open to "any number of approaches to understand the organization's current operating environment." After deliberate reflection and discussion with business professionals and organizational consultants, it was determined that because of the familiarity with and common practice of SWOT analysis in organizations, it would be beneficial to include it as a checkpoint in the OEC.

Major content edits were applied to Step 3 to justify the universal criteria of merit and explain the application of the checkpoints. In addition, a sixth checkpoint was added to Step 3 as a result of Step 5 being eliminated. The additional checkpoint included specific examples to illustrate methods of implementation for the checklist user. Clarifications were made to each checkpoint to address questions and comments such as, "Do all of the measures need to be used in every evaluation?" "You can't always identify measures a priori," and "Maybe say something about the workshop in the checklist item." The overwhelming majority of written comments received for Step 3 were classified as minor content and directed toward improving the clarity and ease of use.

All modifications made to Step 4 were minor content edits that were intended to clarify the three checkpoints included in this step. For example, illustrations were added
to demonstrate the unique aspects of data collection in organizational settings. As noted earlier, Step 5 was eliminated from the original OEC and its nonoverlapping content merged into Step 3. Step 6 in the original OEC was edited to improve clarity and several formatting changes were made to the figures to enhance readability and comprehension.

Although the number and the specific headings of the checkpoints in Step 7 of the original OEC were not modified, explicatory content was greatly expanded in the introduction to this step and, to a lesser degree, in each checkpoint. The inclusion of examples and detailed explanations increased the number of words in this section by a factor of two. Particular attention was given to providing more guidance with respect to communicating evaluation activities and evaluative conclusions to the client. Several paragraphs concerning the inclusion of recommendations in the evaluation report were added to Step 7. Insights into managing client reactions to the evaluative conclusions were added as well.

In summary, the most prominent changes were the expanded content and additional examples incorporated into the OEC. The revisions were intended to enhance the checklist's ease of use and parsimony, as well as build on the strengths identified by the expert panel members. The revised and formatted checklist to be made available for checklist users will focus on a user-friendly format to promote use. For example, all in-line citations and references will be converted to end notes, glossary terms will be highlighted in the text, and icons will be selectively used throughout the document to identify areas of particular import and sources for additional information.
Results of Field Study Using the OEC

An evaluation using the revised OEC to evaluate a for-profit organization's effectiveness was conducted for the second investigation. The purpose of conducting a real-world evaluation was to answer the research question, "How practical is the OEC process for evaluating organizational effectiveness?" The following section presents the findings from the evaluator's and client's perspectives and identifies the revisions incorporated into the OEC based on the field study.

Evaluator Perspectives

The expectation prior to the field evaluation was that the explicit guidelines found in the OEC would facilitate the overall evaluation process to arrive at valid, credible, and useful evaluative conclusions. Overall, this expectation was met. The structured format served as a mnemonic device to focus the evaluation and aided with minimizing diversions during the evaluation process. It also provided a sense of purpose to each activity such that it was possible to see how each element was accretive to the OEC process.

At the initial meeting with the president and a managing partner, the attending managing partner was identified as the primary liaison for ongoing communications and data requests. The information collected to establish the boundaries of the evaluation, conduct the performance needs assessment, and define the criteria of merit was obtained from personal interviews with the managing partners of the firm and document review. The organization engaged an external consulting firm to assist with strategic planning for the organization several years prior to this evaluation. As a result, the company had a well-defined mission statement and an existing SWOT analysis. The SWOT analysis was
very helpful from an external evaluator’s perspective, as it allowed for a quick understanding of the strengths and weaknesses of the organization. Rather than conduct a new SWOT analysis, efforts were made to identify what had changed since the analysis was completed. In this particular case, the analysis was still relevant and only required minor updates to reflect the current situation with respect to market developments, financial health of the firm, and economic environment. Based on this experience, the OEC checkpoint referring to the SWOT analysis was modified to suggest that the evaluator request a copy of the most recent SWOT analysis prior to commencing the SWOT analysis. This revision to the checkpoint supported efforts to increase the parsimony of the OEC.

The universal criteria of merit provided a foundation for establishing the measures and performance standards. Without the explicit criteria of merit included in the OEC, the evaluation may have been severely delayed or indefinitely suspended due to the organization’s emphasis on an informal work environment and limited use of performance monitoring and evaluation, among other contributing factors. The universal criteria of merit stimulated discussion on the contextual matters that define an effective organization, and the performance needs assessment was used to identify contextual criteria of merit.

More guidance than expected was required for the importance weighting task inasmuch as the senior managers rated nearly every criterion as “high” importance, and many were rated as “emergency.” The rating level “emergency” is intended to indicate the temporal urgency of a particular criterion. After discussion with the client, it was determined that none of the criteria was considered urgent and important, and adjustments were made accordingly to the importance ratings. It was also found to be of
value that the senior managers all "publicly" agreed on the importance weightings as well as the criteria, measures, and performance standards. This addendum as well as clarification on the importance rating "emergency" was incorporated into the revised OEC.

Because of the unique service provided by this organization, there are no direct competitors offering the same service, but there are indirect competitors and companies who have similar organizational characteristics. However, these companies serve different markets with different products (e.g., graphic design services) than the evaluation client. This predicament meant that it was not possible to obtain industry or competitive benchmark data that were specific to the client's business activity. However, benchmark data were obtained for service-related organizations whose business models are similar to the evaluation client's model. The OEC handles this situation with reference to the use of external and pan-industry referents to establish performance standards.

Creating performance matrices was found to be the most challenging and time-consuming aspect of the OEC process. It required considerable effort to create the performance matrices and obtain approval from the client. However, this was not unexpected. The engagement of the client in this process is crucial and appeared to have increased the client's interest in the evaluation outcomes and demonstrate transparency of the evaluation process. A recommendation advising an evaluator-facilitated workshop to develop the performance matrices was included in the revised OEC based on these findings and feedback received from an expert panel member.

Communication with the client was done via e-mail, telephone, and face-to-face meetings. E-mails were exchanged on a regular basis every week during the evaluation,
while face-to-face meetings were held biweekly during the early stages of the OEC process (Steps 1 to 3). Once data collection began, it was not necessary to meet with the client until the presentation of the findings.

Overall, the OEC provided a comfortable level of guidance and kept the evaluation on track. The systematic procedures outlined in the OEC offered discipline to the evaluation process and kept the evaluation from extending beyond the defined scope of the evaluation. The foundational checkpoints completed at the onset of the evaluation equipped the evaluator and the client to address challenging aspects, such as defining contextual criteria of merit, establishing measures, and creating performance matrices. Being able to follow the OEC process in a sequential manner helped to overcome obstacles, since the next steps were known in advance. Hence, when a specific issue could be put off until a later date, it was, and allowed the process to move forward.

Client Perspectives

This was the first organizational evaluation that the client had engaged in, excluding financial audits that were conducted by an external public accounting firm. The proposition under consideration was that use of the OEC would facilitate the evaluation process and result in a valid, credible, cost-effective, usable, and ethical evaluation of organizational effectiveness. As was noted in Chapter 4, the OEC would be judged an unacceptable guide to evaluation if it did not focus on the organization’s needs (utility), answer the questions it was intended to address (validity and cost-effectiveness), present defensible conclusions (credibility and validity), and issue findings that were independent of the biases of the evaluation participants (fairness and credibility).
To facilitate discussion regarding the evaluative conclusions, a written evaluation report was provided to the client one week prior to the oral presentation of the results. Immediately following the presentation of the evaluation findings, an interview with two of the managing partners was conducted to assess the evaluation's validity, credibility, utility, cost-effectiveness, and ethicality.

Overall, the client was “very pleased with the evaluation and the process” and noted that the results confirmed what was believed by the client to be deficient areas within the organization. With respect to the evaluation’s validity, the client found favor with the use of “hard” data such as financial ratios and the use of sound techniques to collect data. The inclusion of quantitative data was highly appealing to one of the interview participants, while the other tended to favor “thick” description along the line of storytelling. The OEC includes recommendations for utilizing both quantitative and qualitative measures and multiple data sources to strengthen the validity and credibility of the findings.

When asked specifically about the credibility of the findings, the client was pleasantly surprised at the completeness of the evaluation findings as illustrated by the comment, “I was surprised you picked up as much as you did. I did not think you had so much information to work with. It tells me something about the tool you created—it seemed to work.” In terms of time, effort, and disruption to the organization, the client suggested there was “very low impact on the organization” at an “extremely low cost for what you produced.” This suggests the OEC’s explicit guidance facilitates a cost-effective organizational evaluation with limited disruption to the organization.

With respect to the fairness of the evaluation, the client indicated that the insights uncovered were more than anticipated and that evaluation was complete and
impartial. When asked about the utility of the evaluation, both participants commented on its relevance and indicated the organization would integrate several of the metrics included in the evaluation into their monitoring and performance measurement system. In addition, the recommendations were of considerable interest, and the participants expressed their intent to act on them. It should be noted that all of the recommendations included in the evaluation report were micro-recommendations (i.e., operational recommendations that “fell out” during the evaluation process). For example, during the evaluation process, it was discovered that the average collection period for the organization’s credit sales was 55 days. The industry standard is approximately 40 days, and the organization’s credit terms are Net-15 days. This situation implies that the organization is “financing” its clients’ purchases for an extended period of time, placing greater strain on the organization’s financial health. A recommendation was made to strengthen the organization’s financial controls by actively monitoring key financial indicators and giving consideration to reducing the collection period. Beyond the minimal staff time required to follow-up on past due invoices, there is limited cost to acting on this recommendation, and it is well within reach of the organization. The managing partners agreed. Following the discussion of the recommendations, one of the managing partners commented, “We know these things. But to have somebody tell you makes a big difference.” As a result of the significant interest shown in the recommendations portion of the evaluation, several paragraphs were added to Step 6 of the OEC to address the limitations and consequences of including recommendations in the evaluation report.

A deficiency in the evaluation findings noted by the client was the exclusion of any reference to the organization’s inability to focus its efforts. In other words, the
organization perceives its ability to execute and “stick to” a plan as weak. This weakness should have been discovered during the SWOT analysis. This gap in the evaluation highlights the importance of engaging the client in the review of the draft report as outlined in Step 6 of the OEC.

At the conclusion of the meeting, follow-on support was offered to the client to answer general questions and clarify the implications of specific findings at a future date.

Based on the findings from this case study, several points of clarification were added to the OEC (as outlined in this section). One of the more significant additions pertained to the inclusion of recommendations in the evaluation report. Elements of the OEC that were reinforced by this case study included the need for multiple measures, transparency in the development of the performance matrices, brevity and clarity in reporting, the importance of latent contextual variables (e.g., organizational culture), and the value of identifying a primary liaison early on in the evaluation process.
CHAPTER VI

CONCLUSIONS

This dissertation began with a review of the state of evaluation within organizations and the barriers that have prevented the development of satisfactory approaches to evaluate organizational effectiveness. The argument put forth suggested that the emphasis on the dissimilarities among organizations and their stakeholders has led to a variety of models and theoretical frameworks for evaluating organizational effectiveness. As a result, the organizational effectiveness literature has become noncumulative and fragmented. The OEC framework is an alternative perspective to address the disjointed frames of reference for defining organizational effectiveness with its explicit focus on the common functionality innate to all organizations—regardless of type, size, structure, design, or purpose. The universal criteria of merit checklist and the OEC process feature consistency in level of analysis; resolve problems of overlapping, conflicting, and imprecise criteria of merit; acknowledge that the relative importance of criteria will differ across time and among organizational participants; provide a common theoretical framework based on similarities of organizational function; recognize contextual criteria, yet allow for comparisons to be made across studies; include a specific definition of organizational effectiveness; and delineate criteria for evaluating organizational effectiveness. As such, the tool is intended to provide guidance to
evaluators, organizational consultants, and organizations when designing, performing, or metaevaluating organizational effectiveness evaluations.

The OEC was designed for use in business and industry. Nonetheless, the explicit focus on common organizational activities (input, output, transformation, and exchange) allows the checklist to be used in virtually all organizational settings—regardless of the organization’s type, purpose, structure, or environment. Although the basic framework can be applied to nearly all organizations, the specific measures used may differ according to the organization type (e.g., for-profit or nonprofit), purpose, or other contextual matters. In addition, the inclusion of universal criteria of merit and contextual criteria of merit provide the user with the ability to be situation-specific while working within a general paradigm of organizational effectiveness. This facilitates comparisons across studies, despite potentially varying indicators or measures from organization to organization. From a practical perspective, the OEC is explicitly focused on the input-output cycle of the organization, not on the frame of reference of the actors of the organization under study. This shifts the attention from the value perspective of the researcher or a particular organizational member or actor to value premises based on organizational needs, definition, ethics, and other valid and defensible sources.

Its usage is expected to support the movement toward developing evaluative organizations—organizations that do more than simply perform evaluations. Evaluative organizations encourage and value the evaluative attitude (i.e., thinking evaluatively) and successfully integrate that temperament into the organizational culture. This suggests an evolutionary, yet fundamental, change in the manner in which organizations are viewed, managed, and evaluated. It is hoped that the OEC will augment existing organizational effectiveness evaluation models for a better understanding of the organizational
effectiveness construct and improve organizational performance to benefit organizational members and society.

Implications of the OEC

Implications of the OEC exist in the areas of professional evaluation practice, organizational performance, and theory and research. In this section, the implications in each of these areas are discussed.

Professional Evaluation Practice

As noted above, the OEC is a tool for professional evaluators and organizational consultants to use when designing and conducting an organizational effectiveness evaluation. The pragmatic process provides insight into navigating the organizational evaluation process by framing the evaluation in a manageable set of steps and checkpoints. The approach also expands the evaluator’s domain beyond traditional personnel and training evaluation activities to offer greater utility to the organization.

The systematic process addresses many of the deficiencies found in existing models and is intended to offer a practical method for evaluating organizational effectiveness that provides direction for the evaluator and value to the organization. What’s more, the tool is meant to facilitate the practice of serious evaluation in organizations—advancing from a basic use of skilled evaluation (e.g., quality assurance internal audits) to the internalization of evaluation into the culture of the organization. It is when the evaluative attitude is instilled in the value system of the organization that the organization can be considered an evaluative organization—that is, one that instinctually
reflects on its actions and external environment and continuously improves as a result of those reflections.

To be applied efficiently and parsimoniously, it is assumed the user has developed some level of proficiency in research, measurement, and evaluation—or at least has access to those who do. What’s more, the ability to negotiate the complexities found within organizational settings (e.g., multiple strategic constituencies, time sensitiveness, power and politics, role ambiguities within and across functions, and multiple and conflicting goals) is an important skill for the OEC user. Although the OEC is comprehensive and includes numerous examples illustrating key points, familiarity with organizational settings and behavior will aid the user in conducting a successful evaluation. The implications for evaluator competence should not be understated. Although the OEC clearly promotes a participatory approach to evaluation, the evaluator is responsible for bringing to the evaluation specific competencies with regard to the fundamental evaluative operations (i.e., grading, ranking, scoring, and apportioning), secondary operations (i.e., weighting, barring, and stepping), and an understanding of the complex processes that may contribute to or detract from organizational effectiveness. Furthermore, familiarity with resistance-management strategies (Ridley, 1991) and the identification and alleviation of excessive evaluation anxiety is recommended.

Organizational Performance

Organizations may find value from the OEC in at least two areas: Use of the OEC to facilitate an internal evaluation of organizational effectiveness and as a method for comparing multiple divisions or a consortium of organizations. When used
to conduct an evaluation using internal resources, the internal evaluator (or evaluation team) can rely on the OEC framework to structure and manage the evaluation. Specific steps, checkpoints, or tasks can be assigned to individuals or groups to engage organizational members and reduce the amount of time required to complete the evaluation. In addition, the OEC helps both internal evaluators and executives navigate through an enormously complex open-system environment by focusing on the fundamental aspects of input-transform-output-exchange, highlighting those organizational properties most critical to organizational effectiveness. As a result, the OEC illuminates areas requiring attention that may be overlooked without such a framework. It also ensures that the evaluation focuses on both the short-term and longer-term aspects of organizational effectiveness.

Another implication for organizations is the use of the OEC as a tool for comparing multiple divisions or strategic business units within an organization or for the evaluation of a consortium of independent organizations such as charitable organizations and churches. The systematic approach and universal criteria of merit focusing explicitly on the organization's ability to survive and maximize returns provide a common framework from which to evaluate organizational effectiveness. The findings may be used to identify areas for improvement, prioritize investments or other forms of funding, and assist decision making regarding the allocation of nonpecuniary resources, among others uses.

Theory and Research

Although the OEC process holds much potential for improving evaluation practice within and of organizations, its contribution to theory is found with its
theoretical framework based on similarities of organizational function. This allows for the generalizability of the model for all types of organizations to be tested, and presumes the definition of organizational effectiveness does not change when considering different types of organizations or different stages of development. It also builds on the current understanding of the organizational effectiveness construct and its properties, rather than purport yet another divergent perspective on assessing organizational effectiveness. The universal criteria of merit, explicit definition of organizational effectiveness, and contextual sensitivity provide a solid foundation upon which to test predictions or hypotheses that are derived from this common theoretical perspective (e.g., organizations that perform well across the dimensions provide value to its participants). In addition, specific organizational actions taken can be evaluated in terms of the extent to which they produce the anticipated effects (e.g., initiatives implemented to improve performance on the criteria “evaluative” and “fiscal health” should improve organizational effectiveness).

Another contribution is with respect to an emerging strand in the evaluation discipline that focuses specifically on organizational evaluation and the development of the evaluative organization. The deep-rooted integration of organizations in modern society require well-functioning organizations that maximize effectiveness and minimize harm from their activities to benefit all members of society—even those members who are not considered organizational participants in the strict definition of the term. It is hoped that the attempts made in this dissertation to offer a new perspective for evaluating organizations will provoke evaluators and organizational consultants to put the OEC into practice to help organizations identify areas to improve and possibly break the conceptual and empirical impasse toward a common framework for evaluating
organizational effectiveness. As the awareness of serious evaluation is fostered, it is expected that additional theory-building and empirically grounded efforts will emerge to support these initiatives and stimulate further inquiry based on a common framework that cumulatively adds to the understanding of organizational effectiveness.

Limitations of this Research

Although I believe the OEC is a comprehensive, commonsensical, and unique process-oriented checklist for evaluating organizational effectiveness, I am aware that critics will find a number of areas of fault. The diverse perspectives of various organizational researchers will undoubtedly result in disagreement and apprehension with the concepts proposed in this approach. Nevertheless, the examination, critique, and further development of the ideas presented herein are welcomed and encouraged.

A major limitation of the OEC is its status as a “work in progress.” As with all evaluation checklists, a number of years of field-testing and use are required to adequately validate and refine the tool. Repeated use of the OEC is the preferred method for validation. Inasmuch, the single-case field study was an attempt to establish a reasonable level of validity of the new tool. It was felt that some exposure to a real-world setting was better than no exposure at all and that feedback from the evaluation client would provide some insight for improving the OEC process. As a result, issues that may not have been foreseen in a laboratory or behind a desk were more likely to be discovered in a field setting.

A second limitation was the evaluation client selected for the field study. Although specific selection criteria were used to qualify the organization for inclusion in the study, additional aspects of the OEC process may have been illuminated had the organization been multidivisional, larger, at a different developmental stage, or in a
hypercompetitive marketplace, among a variety of other contextual differentiators. In addition, multiple field studies of organizations of varying types, structures, and purposes would have provided exposure to a broader range of organizational characteristics.

On a similar note, the selection of expert panel members was another limitation of this study. The selection was one of convenience, yet purposeful in its attempt to include a diverse array of targeted users and subject-matter experts. However, only one of the expert panel members was specifically engaged in working with nonprofit organizations, and this may have limited the insights regarding the ability to generalize the OEC across organization types. The small sample size and its adverse impact on statistical power also limit the ability to make inferences to a wider audience of targeted users and to use advanced statistical analysis techniques. Nevertheless, the extensive qualitative feedback received from the expert panel members provided detailed and insightful perspectives that served to highlight deficiencies as well as strengths of the OEC.

Future Areas for Exploration

As noted throughout this dissertation, the definition and conceptualization of organizational effectiveness are ambiguous, and few advances have been made that address the fragmented nature of the construct. Consequently, it is hoped that this dissertation has made some headway toward a common framework for conceptualizing and evaluating organizational effectiveness that moves us closer to an accretive approach.

One area of research worth exploring would be linking organizational effectiveness evaluations to organizational survival. In other words, empirically
demonstrating that an organization deemed effective survived and thrived. This investigation could include ascriptive evaluations of organizational effectiveness over a number of decades for firms with continuing operations, as well as those with discontinued operations. This would best be done using a large sample of publicly traded organizations to permit access to historical data in a practical manner. Another approach would be to initiate a longitudinal study that evaluated organizations to gauge their effectiveness, survival, and maximization of returns during the period under study.

Other areas of interest include answering questions such as, “To what extent do environmental and situational characteristics allow us to predict organizational effectiveness?” and “How do effectiveness profiles change over the lifecycle of the organization?” In addition, future areas to explore include inquiry into the development of a universal set of measures that are connected to the universal criteria of merit presented in the OEC. Except for the accounting profession, there is little agreement on what measures best reflect organizational attributes. Standardized measures would facilitate comparisons across organizations and encourage the accumulation of data from single-case studies as well as multiple-case studies.

Evaluation in organizations is not a new endeavor, but it is one that has been limited primarily to procedural compliance, quality assurance, process audits, and product and personnel evaluation. It has not been seriously developed at the organizational level and certainly not to the point where organizations are working toward becoming evaluative organizations. I view this dissertation as a first attempt at defining a new strand of evaluation—one that purposely injects evaluation logic and methodology into the organizational vernacular and serves the needs of organizations as evaluation fulfills its role as a transdiscipline that holds life together.
REFERENCES


Appendix A

Organizational Effectiveness Evaluation Checklist (OEC)
ORGANIZATIONAL EFFECTIVENESS EVALUATION CHECKLIST
Introduction

The Organizational Effectiveness Evaluation Checklist (OEC) is a tool for professional evaluators, organizational consultants, and management practitioners to use when conducting an evaluation of organizational effectiveness. It can also be used for designing and metaevaluating an organizational effectiveness evaluation. Hence, the systematic process provides direction for the evaluator and value to the organization.

The OEC is designed for use in business and industry. However, the explicit focus on common organizational activities (input, output, transformation, and exchange) allows the checklist to be used in virtually all organizational settings regardless of the organization’s type, purpose, structure, or environment.

Although the basic framework can be applied to nearly all organizations, the specific measures used may differ according to the organization type (e.g., for-profit or nonprofit), purpose, and other contextual matters. The contingency aspect of the OEC approach permits this flexibility to account for changes in environmental conditions and shifts in organizational direction.

Other characteristics of the OEC include its temporal relevance (i.e., balancing short-run considerations with long-run interests), ability to deal with conflicting criteria, practical versus theoretical significance, modest usage of organizational goals, and generalizability of the approach.

Throughout this checklist the term organization refers to planned units, deliberately structured for attaining specific goals or purposes. Organizations are viewed as open-systems that transform inputs into outputs and exchange the outputs with the environment for something in return. Organizational effectiveness is defined as the extent to which the organization provides sustainable value through the purposeful transformation of inputs and exchange of outputs, while minimizing harm from its actions.

The first part of the checklist provides a process overview followed by an outline of the checkpoints for the user to ensure that important components are not omitted in error or oversight. The second part of the checklist provides explanations of each step and checkpoint, and includes the universal criteria of merit for evaluating organizational effectiveness. A glossary of important terminology used throughout the document is provided at the end of the checklist as a quick reference guide for the user. The glossary terms are CAPITALIZED throughout the checklist. In addition, guiding questions are included in the sidebar to stimulate use and provide direction to the user.
Process Overview

The OEC has six steps: (1) establish the boundaries of the evaluation; (2) conduct a performance needs assessment; (3) define the criteria of merit; (4) plan and implement the evaluation; (5) synthesize performance data with values; and (6) communicate and report evaluation findings. Within each step are a series of checkpoints intended to guide the evaluator. Embedded within this framework are the universal criteria of merit for evaluating organizational effectiveness.

The OEC is an iterative process. The iterative nature suggests that the user will benefit from going through the checklist several times as new information is discovered or problems identified that may require modifications to reach an appropriate appraisal for each checkpoint. For example, the identification of performance-level needs is a checkpoint under Step 2. The performance-level needs are used to generate contextual criteria (Step 3) for which performance data are collected (Step 4). While collecting performance data, previously unrecognized environmental conditions may be identified and justify the inclusion of additional contextual criteria. These developments should be anticipated, even sought out, to ensure due consideration is given to each conclusion made during the evaluation process.

Although the OEC is iterative, the sequence of the checkpoints is of logistical importance (i.e., criteria must be defined prior to collecting performance data) and ordered for efficiency. It is recommended to make a quick trial run through the OEC prior to committing to the evaluation to ensure an evaluation of the organization’s effectiveness is appropriate based on the questions to be answered and the data that will need to be obtained to make evaluative conclusions.
Outline of the OEC

1. Establish the boundaries of the evaluation.
   - 1.1 Identify the evaluation client, primary liaison, and power brokers.
   - 1.2 Clarify the organizational domain to be evaluated.
   - 1.3 Clarify why the evaluation is being requested.
   - 1.4 Clarify the timeframe to be employed.
   - 1.5 Clarify the resources available for the evaluation.
   - 1.6 Identify the primary beneficiaries and organizational participants.
   - 1.7 Conduct an evaluability assessment.

2. Conduct a performance needs assessment.
   - 2.1 Clarify the purpose of the organization.
   - 2.2 Assess internal knowledge needs.
   - 2.3 Scan the external environment.
   - 2.4 Conduct a strength, weakness, opportunity, and threat analysis.
   - 2.5 Identify the performance-level needs of the organization.

3. Define the criteria to be used for the evaluation.
   - 3.1 Review the universal criteria of merit for organizational effectiveness.
   - 3.2 Add contextual criteria identified in the performance needs assessment.
   - 3.3 Determine the importance ratings for each criterion.
   - 3.4 Identify performance measures for each criterion.
   - 3.5 Identify performance standards for each criterion.
   - 3.6 Create performance matrices for each criterion.

4. Plan and implement the evaluation.
   - 4.1 Identify data sources.
   - 4.2 Identify data collection methods.
   - 4.3 Collect and analyze data.

5. Synthesize performance data with values.
   - 5.1 Create a performance profile for each criterion.
   - 5.2 Create a profile of organizational effectiveness.
   - 5.3 Identify organizational strengths and weaknesses.

6. Communicate and report evaluation activities.
   - 6.1 Distribute regular communications about the evaluation progress.
   - 6.2 Deliver a draft written report to client for review and comment.
   - 6.3 Edit report to include points of clarification or reaction statements.
   - 6.4 Present written and oral reports to client.
   - 6.5 Provide follow-on support as requested by client.
Step 1: Establish the Boundaries of the Evaluation

Establishing the boundary of the evaluation explicitly defines what is and is not included in the evaluation. The complexity of an organization, its multiple constituencies and perspectives, and the open-system environment require that the effectiveness construct be bounded at the outset of the evaluation. The checkpoints included in this step address specific issues that allow the scope of the assessment to be defined and appropriately circumscribed.

1.1 Identify the evaluation client, primary liaison, and power brokers.

The evaluation client is generally the person who officially requests and authorizes payment for the evaluation. It is also the person to whom the evaluator should report. However, in some cases, the client may consist of a committee or governance board. In such cases, it is important that a chairperson or group leader be identified to promote direct communication and accountability. The primary liaison is generally the primary source of information with respect to the evaluation's purpose and timeframe and is the designated interface between the evaluator and the organization. This individual is also the "go-to" person when issues arise related to data or personnel accessibility. In some organizations, an evaluation may be contracted to establish a case for or against power brokers in the environment. Because of this situation, it is important to identify the person or persons in a position of power to leverage or utilize the evaluative conclusions.

1.2 Clarify the organizational domain to be evaluated.

The organizational domain refers to specific attributes associated with the organization. The domain is typically circumscribed by the constituencies served, technology employed, and outputs (i.e., goods or services) produced. The organization's primary activities, competencies, and external forces influence and shape the organizational domain. In other words, the organizational domain is defined by the operating environment with special focus on identifying customers or clients served (e.g., the target markets), products and services offered, and competitive advantages of the organization. In a loosely coupled organization (e.g., a holding company or university), there may be several domains in which the organization competes or specializes. The same condition can be found in multidivisional organizations, where the domain under investigation may be a strategic business unit or a stand-alone operating division within a business unit. Identifying the organizational domain to be evaluated is an important consideration to focus the evaluation, as well as to ensure the evaluation is directed toward the proper domain to avoid misleading evaluative conclusions about the organization's effectiveness.
1.3 Clarify why the evaluation is being requested.

The purpose of an organizational effectiveness evaluation affects the type of data required, data sources, degree of EVALUATION ANXIETY present or that could develop, amount of cooperation or collaboration required of the ORGANIZATIONAL PARTICIPANTS, and overall assessment strategy, among other factors. A FORMATIVE EVALUATION to identify areas for improvement will be quite different from an evaluation to determine which division to close or agency to fund. Some reasons why an evaluation of organizational effectiveness may be requested are to identify areas for improvement in a particular business unit; to facilitate prioritization of strategic initiatives regarding organizational performance; to assist decision making regarding the allocation of resources; to improve the value the organization delivers to its PRIMARY BENEFICIARIES; and to strengthen an organization's competitive position in the marketplace.

1.4 Clarify the timeframe to be employed.

Judgments of effectiveness are always made with some timeframe in mind. It is important to clarify over which period effectiveness is to be evaluated. Long-term effectiveness may look much different when short-term effectiveness criteria (e.g., monthly profitability) are used.

It is most practical to consider a timeframe of one year to five years. Anything less than one year may not fully reflect the contribution of various strategies and initiatives that require some period of maturation to show effect. Periods longer than five years may become irrelevant due to the complex and generally turbulent environment in which most organizations operate. For example, when the organization competes in a high-tech environment with product lifecycles of six months from introduction to obsolescence, a timeframe of more than two years becomes immaterial to the assessment of the organization's ability to maximize returns and survive in the rapidly changing environment.

For most organizations, a three-year perspective is long enough to allow strategic initiatives to take effect and trends to be identified, yet short enough to maintain relevance to the operating environment. When conducting an ASCRIPTIVE EVALUATION of organizational effectiveness, the timeframes noted above do not apply.
1.5 Clarify the resources available for the evaluation.

The common resources required for all evaluations include money, people, and time. It is important to clarify the financial resources (i.e., budget) available to conduct the evaluation, who will be available as key informants, who will facilitate the data collection, who will approve the evaluator's request for data, and the time available for the evaluation. When an external evaluator conducts the evaluation, these items are generally clarified in the evaluation proposal or soon after its acceptance. However, it is also important for internal evaluators to clarify the resource issue upfront to avoid running out of time or funding and ensure access to personnel and data. Moreover, it is important to identify which organizational resources are available for the evaluator's use, who will provide permission to collect data, and who will confirm the authorization for doing so.

When members of the organization play a role in data collection or substantively participate in other aspects of the evaluation, the evaluator may have to factor in a longer timeframe to complete some activities since these individuals (i.e., collaboration members) may be required to perform the new evaluation assignments in addition to their current duties. What's more, collaboration members may require special training to perform the tasks adequately. A systematic and thoughtful approach should be used for the identification of collaboration members, scheduling critical evaluation activities, clarifying the roles of the evaluator and collaboration members, and other aspects associated with performing a collaborative evaluation.

1.6 Identify the primary beneficiaries and organizational participants.

Organizations operate in dynamic and complex environments that require the consideration of contextual criteria of merit for the specific organization being evaluated. Based on this requirement, the dominant coalition or primary beneficiaries are engaged to facilitate the performance needs assessment (Step 2), identify contextual evaluative criteria, and determine importance weightings of the criteria of merit (Step 3).

The identification of the primary beneficiaries serves an important function with respect to contextual criteria of merit. Although there are a number of persons and groups that benefit directly and indirectly as a result of organizational activities, the primary beneficiaries are stakeholders that are uniquely served by the outcomes of the organization's activities. In the case of most businesses, the organization was created to generate wealth for its owners. Hence, owners would be the primary beneficiary. For a nonprofit organization (i.e., churches, schools, hospitals), the primary beneficiaries are the intended recipients of the service. For commonwealth or government organizations, the public is the intended primary beneficiary.
In organizations where the primary beneficiaries are not in a position to be knowledgeable of organizational functions and constraints (e.g., students in an elementary school), the stakeholder focus shifts from primary beneficiaries to the **dominant coalition**. The dominant coalition, sometimes referred to as the power center, includes organizational participants (e.g., parents of elementary school students) that direct the organization in its concerted efforts to achieve specific objectives. In essence, this group's support or lack thereof impacts the survival of the organization.  

Beyond the primary beneficiaries and dominant coalition, a broader group of organizational participants should be identified to aid in identifying **side impacts**, both positive and negative, that affect nonintended audiences. Organizational participants can be considered from two perspectives: (1) those persons who act on behalf of the organization and (2) those who are external to the organization acting on their own behalf and either affect members' actions or are affected by them. Those persons who act legally on behalf of the organization are referred to as organizational members and include employees, management, advisors, agents, and members of governance boards, among others. The organizational participants external to the organization—for example, shareholders, customers, vendors, and government agencies, among others—are referred to as organizational actors. The identification of organizational participants leads to the identification of those persons who are affected by organizational activities and have a stake in the organization's survival and maximization of returns.

1.7 **Conduct an evaluability assessment.**

The information collected to this point allows the evaluator to perform an **evaluability assessment** prior to committing significant resources to an **evaluand**. Evaluability assessment is the determination of the appropriateness of conducting an evaluation. It is used to determine if the evaluand is "ready" to be evaluated based on the existence of goals, resources, data accessibility, and how the evaluation is intended to be used. It is useful in situations where goals of the organization are known, but the measures of the goals are not yet defined. What's more, the evaluability assessment allows for a preliminary review of the appropriateness of the organization's goals. Although the evaluator may not have specific knowledge about the organization's resources or capabilities at this point in the evaluation, the organization's goals can be assessed in terms of the degree to which they are specific, measurable, attainable, realistic, and time-based. In addition, official goals can be restated as operational goals and the distinction between goals and constraints can be made clear.

Conventional evaluability assessment considers four questions: (1) What are the goals of the organization? (2) Are the goals plausible? (3) What measures are needed and which are available? (4) How will the evaluation be

---

It is recommended to make a quick trial run through the OEC prior to committing to the project to ensure an evaluation is appropriate based on the questions to be answered and the data that will need to be obtained to make evaluative conclusions.

---

Are the organization's goals specific, measurable, and realistic?

What performance data is (or could be) made available?

Are the available resources adequate for the purpose of the evaluation?

Is the evaluation appropriate considering the developmental stage of the organization?
Since everything is evaluable, to some extent in some contexts, the issue of evaluability is a matter of degree, resources, and circumstance, not of absolute possibility. Hence, while everything is evaluable, not everything is evaluable to a reasonable degree of confidence, with the available resources, in every context, for example, the atomic power plant program for Iran in April 2006, when access was denied to the U.N. inspectors.  

Step 2: Conduct a Performance Needs Assessment

The performance needs assessment provides background information about the organization and its operating environment. The systematic approach results in the identification of performance-level needs to provide the evaluator with insight to guide the evaluation, set priorities, and develop contextual criteria to supplement the universal criteria of merit. In contrast to wants or ideals, needs are things that are essential for organizations to exist and perform satisfactorily in a given context. As such, a performance-level need, sometimes referred to as a fundamental need, is something without which dysfunction occurs. This perspective goes beyond the discrepancy definition where needs are defined as the gap between the actual and the ideal, as it considers a greater distinction between different types of needs (e.g., met, unmet, conscious, and unconscious needs).

The performance needs assessment explores and defines the current state of the organization from internal and external perspectives. It considers the organization’s structure, strengths and weaknesses, opportunities available, and constraints that limit or threaten the organization’s survival or its maximization of return. In addition to providing the required information for identification of performance-level needs, the benefits of conducting a needs assessment include building relationships among those who have a stake in the situation, clarifying problems or opportunities, providing baseline performance data, and setting priorities for decision making.

2.1 Clarify the purpose of the organization.

This checkpoint is used to identify contextual aspects that may be unique to the organization or require additional inquiry. The purpose of the organization is its raison d’être—its reason for existence. The purpose is usually, although not always, found in the organization’s mission and vision statements. In cases

---

Why does the organization exist?

What is the organization’s vision and mission?

Have operative goals been identified?

Are the goals or purposes aligned within the organization?
where a written mission or vision statement does not exist, the organization's purpose can be identified by interviewing senior management and reviewing strategic plans or other organization-specific documents such as departmental plans and objectives, as well as the organization's website.

The mission of an organization is generally action-oriented and promotes the OFFICIAL GOALS of the organization. Official goals are also referred to as public goals and may not be the actual (operative) goals of the organization. Care should be taken to distinguish between official goals and OPERATIVE GOALS, the latter being more relevant when assessing organizational effectiveness.

The vision statement provides a sense of direction, is aspirational, and indicates where the organization wants to be in the future. It commonly includes the organization's purpose as well as its values. A vision statement expresses the end, while the mission expresses the means toward reaching the ends. In organizations where divisional-level or program-level goals do not align with those of the larger organization, serious attention should be given to clarifying the "true" purpose of the organization to avoid misleading evaluative conclusions about the organization's effectiveness.

2.2 Assess internal knowledge needs.

This checkpoint is intended to provide a "pulse check" to get an indication of the health of the organization in terms of knowledge management practices. The output of this checkpoint is documentation of (1) the existing knowledge management practices; (2) knowledge that is required, whether in existence or not, to maximize returns and support long-term sustainability of the organization; and (3) specific knowledge-management needs based on the discrepancy between the actual and required states. As a component of the performance needs assessment, this checkpoint provides another perspective on the organization's needs, with particular emphasis on knowledge management.

The first element of the knowledge needs assessment attempts to define the deliberate knowledge management efforts that are in place. Evidence for these efforts can be found in process flow diagrams, standard operating procedures, training programs, role-playing (i.e., rehearsal) activities, decision-making drills, and collaboration systems, among other sources. The second element considers what information is required to make decisions. This information may already exist as a part of the organization's current knowledge management practices or it may be desired to facilitate decision making activities. Asking questions such as, "What information would make your job easier?" or "How does the organization solve problems?" may reveal unmet knowledge needs within the organization. The third element considers the discrepancies between the current knowledge management practices and those required, but currently not available, to avoid dysfunction and enable the
organization to maximize returns and sustain itself. Based on the findings of the knowledge needs assessment, contextual criteria of merit may be developed for inclusion in the evaluation.7

2.3 Scan the external environment.

The open-system nature of the organization recognizes that external factors affect organizational performance. The purpose of this checkpoint is to identify those factors that constrain or enhance organizational effectiveness and to determine their implications. Environmental scanning is a systematic approach to detecting scientific, technical, economic, social, and political trends and events important to the organization. Porter’s (1980) FIVE FORCES MODEL provides a general framework for performing the environmental scan in organizations.8 The five forces include the intensity of competitive rivalry, threat of substitutes, buyer power, supplier power, and barriers to entry.

In addition to these, it is important to consider other environmental factors including sociocultural, political, legislative, and regulatory forces. Special attention should be given to the organization’s primary competitors, as this group is purposely acting to displace the organization’s position in the marketplace. Effective environmental scanning will enable the evaluator to anticipate changes emerging in the organization’s external environment. The consequences of this activity include fostering an understanding of the effects of external change on the organization and aiding in identifying contextual factors influencing organizational performance.

2.4 Conduct a strength, weakness, opportunity, and threat (SWOT) analysis.

The SWOT analysis focuses on the organization’s internal and external environments. The external environment considered in the SWOT analysis is focused on specific elements that are identified as known opportunities and threats rather than scanning for broad environmental trends (outlined in the previous checkpoint). The purpose of a SWOT analysis is to systematically identify areas where the organization excels (i.e., strengths), areas of weakness, opportunities to leverage, and threats to mitigate. It is important to recognize that strengths and weaknesses are strictly internal to the organization. In other words, they are under the control of the organization. Opportunities and threats, in contrast to strengths and weaknesses, are external factors beyond the organization’s control. Some opportunities and threats may “fall out” of the environmental scan completed in the previous checkpoint. However, attention should be given to searching for specific opportunities and threats whose impacts on the organization are imminent.
Because this form of analysis is commonly used for strategic planning purposes in organizational settings, an existing SWOT analysis may be available to the evaluator. If this is the case, the current or past analyses should be carefully reviewed to acquire greater insight into the essence of the organization. An existing SWOT analysis does not relieve the evaluator of performing an updated analysis, but does facilitate the present analysis by providing additional perspective. The results of the SWOT analysis facilitate the identification of performance-level needs referred to in the next checkpoint.

2.5 Identify the performance-level needs of the organization.

Each of the preceding checkpoints provides the necessary information to identify the organization's performance-level needs. Performance-level needs are needs that, if not met, lead to dysfunction within the organization. Identifying performance-level needs reveals contextual evaluative criteria that can be added to the universal criteria of merit for evaluating organizational effectiveness. Examples of a performance-level need and a tactical-level need (i.e., an instrumental need that addresses the performance need) are shown below.

**Performance-level need:** Strategic initiatives within the organization need to be aligned.

**Tactical-level need:** Foster awareness and internalization of the organization's strategic objectives with each employee.

Step 3: Define the Criteria for the Evaluation

The OEC includes two categories of criteria: universal criteria of merit that apply to all organizations and contextual criteria based on a performance needs assessment or other relevant values (e.g., operative goals). Criteria are necessary for credible evaluative conclusions to be made about an organization's effectiveness. Guidelines for a list of criteria of merit include:

1. The list should refer to criteria and not mere indicators.
2. The list should be complete (i.e., no significant omissions).
3. The list should be concise.
4. The criteria should be nonoverlapping.

5. The criteria should be commensurable.

6. The criteria should be clear.

7. The criteria should be confirmable.

8. The criteria should be limited to the organizational level of analysis.

9. The criteria should reflect the relation between the organization and its environment.

10. The criteria should allow for the uniqueness of the organization.

11. The criteria should include both the means and ends of organizational activity.

12. The criteria should be stable yet provide the necessary latitude for organizational change and variability over time.

The starting point for any list of criteria of merit is an understanding of the nature of the evaluand and the properties that make it good, valuable, or significant. In the case where the organization is the evaluand and the evaluation pertains to its effectiveness, the question to answer becomes, "What properties are included in the concept of organizational effectiveness?"

The OEC includes 12 universal criteria of merit for evaluating organizational effectiveness. The criteria are grouped into four dimensions to illustrate the connection with the definition of organizational effectiveness. The dimensions include (1) PURPOSEFUL, (2) ADAPTABLE, (3) SUSTAINABLE, and (4) HARM MINIMIZATION. The 12 universal criteria of merit used in the OEC are shown in Table 1 along with suggested measures for each criterion. The specific measures listed are not required to be used; rather, they are included to provide suggestions for measures to consider. Definitions of the criteria are included in the glossary at the end of this checklist.
Table 1: Universal criteria of merit and suggested measures.

<table>
<thead>
<tr>
<th>Dimension</th>
<th>Criterion</th>
<th>Suggested Measures</th>
</tr>
</thead>
</table>
| **Purposeful** | Efficiency | Revenue per employee-hour
| | | Profit per employee-hour
| | | Profit per square foot
| | | Cost per client served
| | | Cost per unit of output
| | | Fixed asset utilization rate
| | | Revenue multiplier
| | | Unit volume per employee-hour
| | | Unit volume per machine-hour
| | | Gross output per employee-hour
| | | Gross output per machine-hour
| | | No. of clients served per employee-hour
| | | No. of billable hours per employee-hour
| | | Gross payroll power
| Productivity | | Planning and goal setting
| | | Extent of routinization
| | | No. of layoffs during the period
| | | Extent of job rotations
| | | Safeguarding assets
| | | Alignment of strategy, mission, vision
| | | Compliance with established procedures
| **Adaptable** | Innovation | R&D expenses as a percentage of net revenue
| | | Training as a percentage of net revenue
| | | New product development rate
| | | No. of new markets entered during the period
| | | Willingness to innovate
| | | Operational process change frequency
| | | Administrative process change frequency
| Growth | | Compounded annual growth rate (revenue)
| | | Profit or fund growth during the period
| | | Relative market share change
| | | New customer revenue growth
| | | New market revenue growth
| | | Change in manpower
| | | Net change in assets
| **Evaluative** | | Feedback system utilization
| | | Performance management system utilization
| | | Task force utilization
| | | No. of new initiatives launched
| | | Percent of internally generated business ideas
| | | Percent of externally generated business ideas
| | | Change initiatives launched during the period
<table>
<thead>
<tr>
<th>Dimension</th>
<th>Criterion</th>
<th>Suggested measures</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sustainable</td>
<td>Fiscal health</td>
<td>Return on net assets/equity /invested capital</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Net debt position</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Free cash flow</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Liquidity ratios</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Profitability ratios</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Expense ratios</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Fund equity balance</td>
</tr>
<tr>
<td>Output quality</td>
<td>Customer satisfaction / loyalty</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Customer retention rate</td>
<td></td>
</tr>
<tr>
<td></td>
<td>External review / accreditation</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Internal quality measures</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Warranty claims</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Service errors</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Response time</td>
<td></td>
</tr>
<tr>
<td>Information</td>
<td>Role ambiguity</td>
<td></td>
</tr>
<tr>
<td>management</td>
<td>Integrity of information</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Timeliness of information</td>
<td></td>
</tr>
<tr>
<td></td>
<td>No. of staff meetings per month</td>
<td></td>
</tr>
<tr>
<td></td>
<td>No. of company-wide meetings per year</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Perceived adequacy of information available</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Access to procedures, rules, and regulations</td>
<td></td>
</tr>
<tr>
<td>Conflict-cohesion</td>
<td>Work group cohesion</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Employee turnover</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Absenteeism</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Workplace incivility</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Commitment</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Bases of power</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Violence of conflict</td>
<td></td>
</tr>
<tr>
<td>Harm minimization</td>
<td>Intra-organizational</td>
<td>Instances of ethical breach</td>
</tr>
<tr>
<td></td>
<td>Results of ethics audits</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Evidence of workforce training</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Evidence of monitoring systems</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Components of organizational justice</td>
<td></td>
</tr>
<tr>
<td></td>
<td>No. of employee accidents</td>
<td></td>
</tr>
<tr>
<td></td>
<td>External / internal audits</td>
<td></td>
</tr>
<tr>
<td>Extra-organizational</td>
<td>Regulatory compliance</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Ecological footprint change</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Environmental controls and monitoring</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Emissions levels (pollutants, noise)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>External audits</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Contribution to the larger system</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Philanthropic activities</td>
<td></td>
</tr>
</tbody>
</table>

*Note. Additional measures can be found in *Handbook of Organizational Measurement*, J. L. Price (1986).*
The universal criteria of merit capture both the means and ends of organizational activities, allow for the comparative study of organizations or subunits, and recognize the uniqueness of the organization being evaluated. When combined with the contextual criteria of merit, the list provides an adequate inventory of the characteristics and properties of an effective organization.9

In organizations where more than five participants will be involved in defining the criteria, measures, and performance standards, it is best to approach this step using a two-phase process. The first phase consists of using a small working group of two or three key informants to work through each of the checkpoints in Step 3. The evaluator and small working group complete the various checkpoints with the intent of developing a draft set of criteria, measures, and performance standards. This working group also identifies potential challenges or points of concern that may arise when the structured discussion is conducted with the larger group. The draft set of criteria, measures, and performance standards are provided to the primary beneficiaries or dominant coalition members in advance of the evaluator-facilitated workshop that occurs in the second phase of this step.

Phase two consists of two 90-minute workshops intended to engage the primary beneficiaries or dominant coalition in an open, yet structured, dialogue and exchange of ideas regarding organizational effectiveness and its dimensions. The checkpoints are used as the framework for the structured discussion and the draft set of criteria, measures, and importance weightings are adjusted as needed during this first review.

The highly compressed time allotment for this workshop is intended to focus the discussion on the specific issues related to criteria development and requires the small working group to be well prepared prior to this workshop. It also requires that the evaluator “sell” the importance of the criteria of merit checklist in advance, because the immediate value of this exercise may not be seen by all participants. This is particularly true if an internal evaluator is used rather than an external evaluator. The second 90-minute workshop is used to review the performance standards and create performance matrices outlined in the last two checkpoints of Step 3.

3.1 Review universal criteria of merit of organizational effectiveness.

The universal criteria of merit consist of the characteristics that define an effective organization. These characteristics are intended to be applicable to all organizations that are deliberately structured for a specific purpose. The universal criteria of merit are reviewed with the client to ensure each criterion and dimension is understood and to stimulate thinking about potential measures that may be used. Refer to Table 1 for the list of universal criteria of merit and suggested measures for consideration.
3.2 Add contextual criteria identified in the performance needs assessment.

The information collected in the performance needs assessment may have revealed additional evaluative criteria that are unique to the organization. These criteria may result from the political, social, or cultural environment; developmental stage of the organization; current situational issues threatening the survival of the organization; or other matters that are unique to the organization at the particular time of inquiry. When considering contextual criteria of merit in multidivisional organizations, it is important to look across the organization in addition to within the particular unit or division to ensure optimization in one unit does not result in suboptimization in another.

3.3 Determine the importance weightings for each criterion.

The weighting of criteria by relative importance recognizes that some criteria are more important than other criteria. It also allows for a more complex inference. Weighting is particularly important when the evaluative conclusions for each dimension must be synthesized into an overall evaluative conclusion regarding the organization's effectiveness. When using the OEC to conduct a formative evaluation that uses profiling to show how the organization is performing on the various effectiveness dimensions, weighting may be avoided; the client can use the multiple evaluative conclusions to identify and prioritize areas for improvement. However, when conducting a summative evaluation, it is usually necessary to go one step further than profiling and make an overall evaluative conclusion for the benefit of the client and the utility of the evaluation.

There are a number of different strategies for determining importance weightings, including voting by stakeholders or key informants, using expert judgment, and using evidence from a needs assessment, among others. Two alternative methods for determining the relative importance of criteria are offered here.

The first is a qualitative approach that consists of having the primary beneficiaries or dominant coalition agree on each criterion's importance using the categories of low, medium, high, and emergency. The emergency category is used to reflect the temporal importance or urgency of a criterion at a particular moment in time. It can be applied to criteria that may not be of high importance in the long term, but are of critical importance in the immediate term. For example, if an organization is suffering from cash flow or credit restrictions and only has the ability to pay its employees for 60 days, growth and fiscal health criteria may be categorized as "emergency." Unless these areas are addressed with urgency, the organization's professional development programs or launch of a redesigned website become less relevant as the organization's survival is of immediate concern.
The second method is a quantitative approach using the analytic hierarchy process to derive a set of numeric weights from pair-wise comparisons. Each member of the dominant coalition identifies which of two performance criteria he or she believes is more important and then records the magnitude of the selected criterion’s importance over the criterion not selected. This process is repeated until every criterion has been compared to the others. To determine the weightings, an organization-level matrix is created from each participant’s pair-wise comparisons, the comparisons are normalized, and the importance weighting is calculated. This procedure can be done using spreadsheet software with intermediate-level knowledge of how to use the software.

Note that it is important to have all primary beneficiaries or dominant coalition members agree on the importance weightings in front of each other, before any data are collected. This “public” affirmation reduces the likelihood that participants will change their opinion after seeing the performance data or evaluative conclusions.

3.4 Identify performance measures for each criterion.

The performance measures for each criterion are the factual data that will be collected and synthesized with the values (i.e., criteria) to produce the evaluative claims. It is best to use measures that can be observed and are stable and valid. Whenever possible, include several measures for each criterion, preferably from different sources.

Of equal importance, agreement should be reached on precise nature of the measures and data sources. For example, if revenues are a measure, is revenue recognized when the order is booked, invoiced, or money collected? Are revenues determined before or after the impact of sales incentive plans? Is the change in revenue determined year-over-year or quarter-over-quarter?

In situations where performance cannot be observed or directly measured, the inclusion of multiple measures from different sources will increase the validity and credibility of the findings for the particular criterion and contribute differentially with unique information.

3.5 Identify performance standards for each criterion.

Performance standards are the claims against which performance data are compared. In other words, standards are quality categories of increasing merit; in some organizations, they are referred to as benchmarks. In most cases, organizations will have internal quality standards already in place that can be leveraged. However, it is appropriate to use industry or peer-group performance standards in addition to internal standards. In those cases where the peer group performs poorly, the use of best-demonstrated practices in any sector or industry as a referent is recommended.
In the case of harm minimization, this dimension is an absolute measure and does not rely on comparisons to other organizations. In other words, one organization’s illegal activities are not "less illegal" than another organization’s illegal activities. Both have violated the law. A similar argument applies to ethical requirements.

For those criteria whose importance is identified as essential, a bar should be established to indicate the minimum level of acceptable performance. Below the bar, the organization fails on that particular criterion.

There are specific types of bars that can be used for this operation. A SOFT bar, for example, indicates a minimum level of acceptable performance for a particular dimension or subdimension of the evaluand to qualify for entry into a high-rating category. A GLOBAL bar, sometimes referred to as a holistic bar, would be applied to those criteria where performance below a minimum level means the organization is ineffective overall, regardless of exemplary performance on the other dimensions. For instance, if a global bar has been established for specific ethical standards, an organization would be deemed ineffective if violations of ethical standards were discovered—no matter how well it performed on the other criteria of organizational effectiveness.

3.6 Create a performance matrix for each criterion.

A performance matrix is a tool for converting descriptive data into an evaluative description or judgment. It can be used for determining ABSOLUTE MERIT (i.e., grading) or RELATIVE MERIT (i.e., ranking). The most basic performance matrix includes a rating (e.g., excellent, fair, poor) and a description or definition of the rating. An example of a performance matrix to determine absolute merit for a specific criterion is shown in Table 2; a performance matrix to determine relative merit using a different criterion is shown in Table 3.
Table 2: Performance Matrix for Determining Absolute Merit of the Management Support Criterion

<table>
<thead>
<tr>
<th>Rating</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Excellent</td>
<td>The organization requires only limited and infrequent support from its parent company or external consultants. It is capable of meeting its financial and operational goals with internal human resources and knowledge.</td>
</tr>
<tr>
<td>Good</td>
<td>The organization is self-sufficient, but requires some input from its parent company on specific issues on an infrequent basis.</td>
</tr>
<tr>
<td>Acceptable</td>
<td>The organization is self-sufficient, but requires some input from its parent company on specific issues on a regular (monthly) basis.</td>
</tr>
<tr>
<td>Marginal</td>
<td>The organization is self-sufficient but requires some input and assistance from its parent company on a frequent (several times monthly) basis. Financial and operational goals are somewhat of a challenge to meet without support from the parent company.</td>
</tr>
<tr>
<td>Poor</td>
<td>The organization lacks ability to self-manage without serious support from its parent company or external consultants. Few, if any, organizational goals can be achieved without support from its parent company.</td>
</tr>
</tbody>
</table>

Table 3: Performance Matrix for Determining Relative Merit of the Profit Per Employee Criterion

<table>
<thead>
<tr>
<th>Rating</th>
<th>Net Profit Margin</th>
</tr>
</thead>
<tbody>
<tr>
<td>Best</td>
<td>&gt; 95%</td>
</tr>
<tr>
<td>Better</td>
<td>75%-95%</td>
</tr>
<tr>
<td>Typical</td>
<td>50%-74%</td>
</tr>
<tr>
<td>Inferior</td>
<td>25%-49%</td>
</tr>
<tr>
<td>Worst</td>
<td>&lt; 25%</td>
</tr>
</tbody>
</table>

Creating performance matrices for each criterion forces organizational participants to think seriously about how they define performance, quality, or value. Engaging primary beneficiaries or the dominant coalition in this process can result in increased buy-in for the evaluation, generate deeper interest in the evaluation process and outcomes, and increase the transparency of the evaluation. The definition of performance characteristics for each criterion can be accomplished during the second evaluator-facilitated workshop session in which each participant or group of participants is assigned two or three criteria and works independently to develop performance matrices for them. Each participant or participant group then shares the proposed performance matrices with the entire group. Revisions are made, and a final version is accepted by the group. This cycle repeats until a performance matrix has been created for each of the criteria and subcriteria.
It is advisable to fully develop all performance matrices prior to collecting any data to avoid the temptation to manipulate the evaluative descriptions and cut-offs to achieve positive ratings. This is particularly true for internal evaluators who may feel pressure from colleagues or superiors to produce a favorable evaluation.

Step 4: Plan and Implement the Evaluation

Items covered in this step focus on the evaluation plan and implementation. The primary determinants influencing the evaluation plan include evaluation team skills, organizational design and structure, available resources, and intended uses of the evaluation. The evaluation team (or individual) skills may narrow the evaluation design options, unless external resources are subcontracted. The organizational design, structure, and other contextual factors (e.g., culture and developmental stage) will also influence which type of data collection methods and sources are most appropriate. In some cases, participants will be willing and able to support data collection efforts. In most organizational evaluations, data sources are internal (e.g., organization members and archival records). The context will support certain data collection methods and inhibit others.

4.1 Identify data sources.

Considering the measures identified in Step 3, the sources of data can be identified. Although many of the measures call for archival data stored as documents or records, the use of observations, surveys, and interviews with various organizational participants will provide a fuller picture of the organization’s performance and assist in uncovering side effects or side impacts that may not be revealed from archival data. If the organization being evaluated is a subunit in a vertically integrated organization, information from upstream units (i.e., those that provide components or service to the unit being evaluated) and downstream units (i.e., those that receive components or service from the unit being evaluated) will support triangulation of the data.

4.2 Identify data collection methods.

The data collection methods are oftentimes directly influenced by a predetermined budget or a short timeframe for the evaluation. To address these influences, the data collection methods must be consistent with the purpose of the evaluation and needs of the organization, be flexible to take advantage of any data source that is feasible and cost-efficient, provide relevant information, and allow for comparisons from multiple data sources for purposes
of triangulation. Every data collection method features some inherent form of measurement error, and using methods that have different types of bias guards against inaccurate conclusions. In addition, using multiple data collection methods and sources reduces the probability the results are due to artifacts of a given method, but represent a truer measure of organizational effectiveness. Hence, the credibility of the evaluation findings is strengthened when using multiple data collection methods and sources.

4.3 Collect and analyze data.

Data collection activities should be nonintrusive and nondisruptive (to the greatest extent possible), cost-efficient, and feasible given the available resources. Although document and record review are likely to have the lowest impact on the organization, they should be managed to minimize frequency of requests and interruptions to the day-to-day activities of the organization's members. Awareness of how the evaluator may affect the organization should be considered whenever activities require the evaluator to interface with the organizational members or actors. For example, when collecting data from highly skilled technologists or professionals whose billable services can be charged out at $400 and up per hour, a 30-minute interview can carry quite a price—both in pecuniary terms and in opportunity costs. This illustration reinforces the point that care should be taken when developing the data collection methods and instruments.

Once the performance matrices have been defined and the data collected, data analysis begins. It is important to consider the client's and other stakeholders' understanding of and requirement for various types of data analysis. The primary consideration is that the results of the evaluation are clear, comprehensible, and credible to the client. For most organizational evaluations, qualitative and quantitative data will be collected. Regardless of the type of data collected, the analysis should be systematic and rigorous and will most likely include rich description of observed processes or interviews, as well as descriptive statistics that include measures of central tendency, variability, and relationships among variables.

Step 5: Synthesize Performance Data with Values

Synthesis is the combining of factual data and values into an evaluative conclusion. It is the final step in the logic of evaluation and is a primary distinction between evaluation and research. The synthesis process is what allows the evaluator to move beyond simply describing "what's so" to answering the question, "So what?"

When the evaluation is used for formative rather than summative purposes, profiling the performance on the various dimensions of organizational effectiveness
may be all that is required. The profile provides the client a graphical summary of performance and highlights areas to address to improve performance. For summative evaluations, synthesis of the various dimensions into an overall evaluative conclusion is required. In this case, the use of a performance matrix (described in Step 3) to synthesize the dimensions into an overall performance score or rating is recommended.

5.1 Create a performance profile for each criterion.

A performance profile is a graphical illustration of how well the organization performs on each criterion according to the performance matrix. With the criteria listed on the vertical axis and the performance ratings shown on the horizontal axis, a performance bar extends outward to the appropriate rating. An example of a performance profile is shown in Figure 1. Note that the profiles for criteria are presented in a different color or shading from the dimension profile (shown as a solid bar). According to this profile, the organization is “good” on the workgroup cohesion and information management criteria, “acceptable” on the output quality criterion, and “marginal” on the fiscal health criterion.

The rating (e.g., excellent, good, acceptable, marginal, or poor) for each criterion is determined by converting the criterion’s performance measures into a score and then calculating the average score for the criterion. For example, assume the rating excellent = 4.0, good = 3.0, acceptable = 2.0, marginal = 1.0, and poor = 0.0. If the performance on the two measures for the criterion was found to be good (3.0) and marginal (1.0) on another, the average criterion score is 2.33. No weighting is involved in this procedure; however, bars would still be utilized where appropriate. This step is repeated for each criterion until scores have been created for all criteria. To arrive at the dimension profile, the importance weightings are applied to each criterion and the numerical weight-and-sum method is used to produce a dimension score.¹⁵

When conducting a summative evaluation, the final synthesis step applies the same procedure used to arrive at the dimension profile, except in this case the importance weightings of dimensions are used and the numerical weight-and-sum procedure is used to arrive a composite score and overall grade of organizational effectiveness.
5.2 Create a profile of organizational effectiveness.

In both formative and summative evaluations, profiling offers an easy-to-use and easy-to-understand tool for the client. When possible, including performance bars for both the dimensions and the criteria condenses broad insight into a concise view. An example of a profile of organizational effectiveness is shown in Figure 2. In this profile, the organization is good to excellent on the purposeful and sustainable dimensions, but requires attention on the adaptable dimension. Defining and implementing actions that improve performance on the adaptable dimension would increase the organization's overall effectiveness.
5.3 Identify organizational strengths and weaknesses.

Based on the organizational effectiveness profile, areas of strength and areas that require attention can be identified. This checkpoint is not intended to provide recommendations to the client. Rather, it is intended to highlight the organization's performance on the various dimensions. The appropriateness of the goal choices as well as performance against organizational goals identified in Step 1 may also be considered at this checkpoint. Inasmuch, the organizational effectiveness profile is used to support statements made regarding the degree of goal attainment. For example, if an organizational goal was to increase revenues by launching a new line of products in a new market and this goal was not achieved, the weak performance found on the innovation and growth criteria may be used to illustrate why achievement of this goal was inhibited. Based on the findings and evaluative conclusions, the evaluator is in a unique position to report on the organization's progress toward meeting its intended purpose or specific goals, and it is in this step that feedback regarding goal attainment is presented to the client.

In some cases, it may be useful to order the list of strengths and weaknesses so that the strongest or weakest dimensions are listed first. This allows the client to quickly grasp the areas of success and those requiring the most serious attention. When communicating organizational deficiencies, it is recommended to avoid personalization of the findings. This is not to suggest that negative findings should not be presented. Instead, it is to suggest that sensitivity must be shown and due consideration given to the level of evaluation anxiety that may be present. For example, using the phrase, "this is what the data tell us," rather than "this is what we found" depersonalizes the findings and can facilitate constructive discussion.\textsuperscript{16}

Step 6: Communicate and Report Evaluation Activities

Communicating about the evaluation itself and reporting evaluation findings serve important roles within the organizational evaluation. During the evaluation, regular communications keep the client informed as to the status of the evaluation and can reduce negative reactions by the client and other organizational members. It can also provide reassurance that the costs (e.g., time, money, and other resources) being incurred are resulting in progress toward completing the evaluation. Reporting is generally thought of as the "end game" of the evaluation process. It is when evaluative conclusions are presented based on the synthesis of the factual data and values.

By and large, a long, written report will not be used in a business context. The client (or primary liaison) may read the entire report, but few others will. To ensure the findings are appropriately presented to others, the inclusion of a well written
executive summary—one page preferred, two pages maximum—supported by a deck of slides using, for example, Microsoft's PowerPoint® software is recommended. The slides should be numbered so it is evident to reviewers who are only seeing two or three slides, that more are available. From a communication style perspective, always summarize first, and then explain. This is particularly important when communicating to executives. The executive summary should not be used as a "teaser"; rather it should give a summary of the evaluative conclusions and recommendations (if any). This is then followed by supporting evidence and explanations that led to the evaluative conclusions presented. The same approach should be followed in the written report; summary statements begin each section, followed by supporting evidence.

The highly charged nature of some evaluations, particularly those that involve summative decisions related to resource allocation or program continuance, require heightened awareness of the clients' reactions. In some situations, the client may make inappropriate inferences from evaluative conclusions and recommendations. It is worthwhile to call attention to what the evaluative conclusions and recommendations imply—and even more importantly, what they do not imply. For example, if the evaluand was a pilot advertising program for a new line of furniture, and the program was determined to be cost-ineffective, a valid implication is that alternative programs are worth consideration. It does not imply that the entire marketing team is incompetent or that the product line should be abandoned. Clearly stating what is implied and what is not implied facilitates a proper interpretation of the evaluative conclusions and may increase the utilization of the evaluation for the betterment of the organization.

6.1 Distribute regular communications about the evaluation progress.

Communicating about evaluation activities on a regular basis serves two primary purposes: It keeps the need-to-know audience up-to-date on the evaluation's progress and can help generate buy-in. The frequency and content of the communications should be based on the client's preferences; a single approach may not suit all organizational evaluations or all audiences within a particular organization.

In most cases, a one-page summary highlighting recent activities, status, upcoming activities, and potential obstacles can be issued biweekly to the client. The distribution may be via e-mail, uploaded to a secure Internet site for broader access, or both. Note that widely distributing information beyond the client and key stakeholders may be considered a political act and should be avoided. If an update is provided orally, it should be followed by a written summary for the benefit of both the client and evaluator.
6.2 Deliver a draft written report to client for review and comment.

The draft written report is an important element in that it allows for clarifications, objections, and other comments to be received from the client prior to submitting a final report. By offering a draft for review, the evaluator is seeking evaluative feedback. This form of feedback performs two primary functions. First, it engages the client. This engagement encourages ownership of the report by the client and may increase the evaluation's credibility and use. Second, findings that are not clearly communicated, are ambiguous, are erroneous, or need verification, may be discovered and corrected prior to the final written report. It is important to reflect carefully on the suggestions made by the reviewers. Care should be taken to ensure that any changes included in the final report based on reviewer feedback do not distort the findings.

6.3 Edit the report to include points of clarification or reaction statements.

Based on the client's feedback to the draft report, the written report is edited as needed. In all cases, the opportunity for the client or a team member to include a reaction statement in the final report should be offered. Although the inclusion of a reaction statement is more commonly found in politically charged or high-stakes environments, it is important to be sensitive to this issue to encourage inclusion of the voices of those who may be less powerful or who have valid arguments related to the evaluation findings. Note that the evaluator's responsibility is not to solve political issues, but to ensure that viable points are raised.

6.4 Present written and oral reports to client.

The final written report should be delivered to the client and a time scheduled to present the findings and evaluative conclusions. The focus of this checkpoint is on brevity and quality of content, rather than quantity. Moreover, the reporting format and delivery method should be chosen to maximize access to the findings (as appropriate), allow for client engagement, and be tailored to needs of various stakeholder groups. The use of quotes from interviews and tables, charts, and other visual tools help encourage audience interaction and support the verbal presentation. When communicating negative findings, it is important to stress the opportunity for organizational learning and improvement while minimizing "finger pointing" toward particular individuals or groups within the organization.

The inclusion of recommendations as part of the evaluation is generally expected in organizational settings. However, care should be taken to limit the recommendations to operational recommendations that "fall out" of the evaluation and can be implemented with little or no extra cost to the client.
These types of recommendations focus on the internal workings of the evaluand and are intended to facilitate improvement efforts. On the other hand, recommendations concerning the disposition of the evaluand (e.g., redirect resources from one business unit to another) are, in nearly all cases, inappropriate due to the evaluator's limited knowledge of the decision space. In those situations where the evaluator does have the required expertise and knowledge to make macro-recommendations, it should be made clear that a different type of evaluation is required—one that assesses alternative options for decision making.18

6.5 Provide follow-on support as requested by client.

Follow-on activities may include answering questions that arise after the client has taken time to absorb the findings, which may be some weeks or months following the delivery of the evaluation report. The evaluator should be available to answer questions, both technical and general in nature, to facilitate utilization of the evaluation.

Including a half-day follow-on session in the evaluation proposal and budget will allow the evaluator to further his or her utility to the client while not detracting from work on other funded projects. In addition, this follow-on support may offer the opportunity to incorporate ongoing performance monitoring and evaluation as a regular activity within the organization for continuous improvement and increased organizational effectiveness. From a project management perspective, it is important to ensure there is a sign-off from the client that the evaluation is complete so that follow-on work is explicitly covered by a different contract, in the case of an external evaluator, or within a new project scope, in the case of an internal evaluator.

Conclusion

The OEC outlines a process framework to evaluate the unique expressions of organizational effectiveness. The distinction of this model from other models for evaluating organizational effectiveness lies in its explicit reliance on the definitional premise of organizational effectiveness based on the input-output cycle inherent in all organizations. Rather than basing the conceptualization of organizational effectiveness on the differences among various organization types, the OEC builds on the similarities of organizations. The practical relevance of the OEC and its ease of use allow organizations to better understand which aspects of the organization need attention to ensure survival and viability of the organization. Users and reviewers of the OEC are encouraged to send criticisms and suggestions to the author at info@EvaluativeOrganization.com.
Glossary of Terms

For a comprehensive review and expanded definitions of evaluation-specific terms, refer to the Evaluation Thesaurus (Scriven, 1991).

**ABSOLUTE MERIT** The unconditional intrinsic value of something that is not relative to another claim. See also relative merit.

**ADAPTABLE** The ability of an organization to change its processes in response to or in anticipation of environmental changes.

**ASCRIPITIVE EVALUATION** An evaluation done retrospectively, generally for documentation or for interest, rather than to support any decision. See also formative evaluation and summative evaluation.

**BAR** A hurdle that sets the minimum standard or acceptable level of performance.

**BENCHMARK** A standard against which performance can be measured.

**COLLABORATIVE EVALUATION** A form of participatory evaluation that involves stakeholders taking on a specific element or multiple assignments related to the evaluation. In all cases of collaborative evaluation there is a significant degree of collaboration between the evaluator and organization participants.

**CONFLICT-COHESION** The cohesion of and by an organization in which members work well together, communicate fully and openly, and coordinate their work efforts. At the other end lies the organization's verbal and physical clashes, poor coordination, and ineffective dedication.

**CRITERIA OF MERIT** Aspects of the evaluand that define whether it is good or bad, valuable or not valuable. Also referred to as dimensions of merit.

**DOMINANT COALITION** A representation or cross-section of horizontal and vertical constituencies within an organization with different and possibly competing expectations. It is the group of persons with the greatest influence on the input-transformation-output process and the identification of means to achieve the agreed upon goal states. See also primary beneficiaries.

**EFFECTIVE** Producing or capable of producing an intended result. See also organizational effectiveness.

**EFFICIENCY** A ratio that reflects the comparison of some aspect of unit performance to the costs (time, money, space) incurred for that performance. It is often used to measure aspects of a process other than just physical output.

**EVALUABILITY ASSESSMENT** The determination of the appropriateness of conducting an evaluation. It is used to understand if the evaluand is "ready" to be evaluated based on the existence of goals, data accessibility, and how the evaluation is intended to be used.

**EVALUAND** The item being evaluated. In the case of the OEC, the organization is the evaluand.

**EVALUATION** The determination of the merit, worth, or significance of something (i.e., the evaluand).
EVALUATION ANXIETY  An abnormal and overwhelming sense of apprehension and fear provoked by the imagined possibility, immanency, or in-process evaluation.

EXTERNAL EVALUATOR  An individual performing an evaluation who is not employed by or affiliated with the organization or program being evaluated. See also \textit{internal evaluator}.

EXTRA-ORGANIZATIONAL  Those items that pertain to matters external to the organization.

EVALUATIVE  The extent to which an organization actively seeks out opportunities for improvement and incorporates the findings via feedback into its planning and operation processes to adapt to the internal and external environment.

FIVE FORCES MODEL  A framework for industry analysis and business strategy development. The five forces include competitive rivalry, bargaining power of suppliers, bargaining power of customers, threats of new entrants, and threats from substitute products.

FISCAL HEALTH  The financial viability of an organization as represented by its financial statements (e.g., balance sheet, income statement, cash flow statement). In most cases, ratios are used to allow for comparison with relevant organizations (e.g., direct competitors or alternatives operating in the same environment).

FORMATIVE EVALUATION  An evaluation done with the intent to improve the evaluand. See also \textit{descriptive evaluation} and \textit{summative evaluation}.

GLOBAL BAR  An overall passing requirement for an evaluand as a whole. Failure on a global bar results in the entire evaluand failing. Also referred to as a hard bar.

GRADING  The assignment of evaluands, dimensions, or subdimensions into a set of named categories. Also referred to as rating.

GROWTH  The ability of an organization to import more resources than it consumes in order to maintain itself.

HARM MINIMIZATION  The extent to which an organization minimizes the negative outcomes created by its activities.

INDICATOR  A factor, variable, or observation that is empirically connected to the criterion variable.

INFORMATION MANAGEMENT  The completeness, efficiency, and accuracy in analysis and distribution of information. This includes cross-level collaboration, participative decision-making, accessibility to influence, and communications.

INNOVATION  The degree to which changes (either temporary or permanent) in process, procedures, or products are intentionally implemented in response to environmental changes.

INTERNAL EVALUATOR  An individual who performs evaluations for and within their organization of employment. See also \textit{external evaluator}.

INTRA-ORGANIZATIONAL  Those items that pertain to matters internal to the organization.

KEY INFORMANT  An individual with specialized skills or knowledge who can provide information or access to information pertaining to a specific topic.
OFFICIAL GOALS Public goals that are stated as the primary objectives of the organization. Official goals may or may not be the actual goals that are used to direct individual behavior.

OPERATIVE GOALS Action-oriented objectives that are intended to direct individual behavior. This type of goal is considered the "real" goal that members of the organization are working toward achieving. It is this type of goal against which an organization's performance may be assessed.

ORGANIZATION A planned social unit deliberately structured for the purpose of attaining specific goals.

ORGANIZATIONAL ACTORS Those persons who are external to an organization, act on their own behalf, and either affect organization member's actions or are affected by them. Examples of organization actors include shareholders, customers, vendors, and government agencies among others. See also organizational members.

ORGANIZATIONAL EFFECTIVENESS The extent to which the organization provides sustainable value through the purposeful transformation of inputs and exchange of outputs, while minimizing harm from its actions.

ORGANIZATIONAL MEMBERS Those persons who act legally on behalf of the organization including, for example, employees, managers, agents, advisors, and members of governance boards. See also organizational actors.

ORGANIZATIONAL PARTICIPANTS Individuals or groups who have a stake in the organization's activities and outcomes. Also referred to as stakeholders. See also organizational actors and organizational members.

OUTPUT QUALITY The quality of the primary service or product provided by the organization may take many operational forms, which are largely determined by the kind of product or service provided by the organization.

PERFORMANCE-LEVEL NEED Anything that is essential to maintain a satisfactory level of performance or state of existence. More specifically, performance needs include met and unmet needs, and conscious and unconscious needs. This perspective goes beyond the discrepancy definition where needs are defined as the gap between the actual and the ideal as it considers a greater distinction between different types of needs. See also tactical-level need.

PRIMARY BENEFICIARIES Those persons for which the organization was created to advantage. For businesses, this would be its owners. For nonprofit organizations (e.g., schools), this would be the program recipients (e.g., students). See also dominant coalition.

PRODUCTIVITY The ratio of physical output to input in an organization. Output consists of the goods or service that the organization provides, while the inputs include resources such as labor, equipment, land, facilities, etc. The more output produced relative to the input, the greater the productivity.

PROFILING A form of grading where performance is indicated for each dimension or component of the evaluand. Profiling does not require differential weightings, and is most useful in formative evaluations.

PURPOSEFUL Acting with intention or purpose.

RANKING The ordering of things from highest to lowest or best to worst.
RELATIVE MERIT The intrinsic value of something compared to one or more alternatives. See also absolute merit.

SIDE EFFECT An unintended effect of the evaluand on the target population.

SIDE IMPACT An unintended effect of the evaluand on the nontarget population.

SOFT BAR A minimum level of performance for a particular dimension or subdimension of the evaluand to qualify for entry into a high-rating category. Failure on a soft bar does not result in failing the entire evaluand. Rather it only limits the highest rating allowed.

STABILITY The maintenance of structure, function and resources through time and more particularly, through periods of stress. This criterion is at the opposite end of the continuum of the innovation criterion.

STAKEHOLDER An individual who has a relevant interest or stake in the organization. This includes persons both internal and external to an organization. See also organizational actors and organizational members.

SUMMATIVE EVALUATION An evaluation done with the intent to aid in decision-making or to report on the evaluand. See also ascriptive evaluation and formative evaluation.

SUSTAINABLE Able to be continued indefinitely.

TACTICAL-LEVEL NEED An action, intervention, or treatment that is intended to address a performance-level need. For example, the need to cost-effectively communicate with organizational constituents (performance need) may require email service or an Intranet web site (tactical needs). More than one tactical need may be identified to address a single performance need. In some settings this type of need is referred to as a treatment-level or instrumental need. See also performance-level need.

Endnotes


7 Special thanks to Dr. Thomas Ward, II for his input on this checkpoint.


16 Comments from Dr. Sharon May were of great help toward clarifying this point.


Appendix B

Critical Feedback Survey Instrument
Critical Feedback Survey

For each of the following statements about the organizational effectiveness evaluation checklist, please circle the number that best reflects the extent to which you agree or disagree. Circle only one rating for each statement.

Respondent Code: xxxx

<table>
<thead>
<tr>
<th>The organizational effectiveness evaluation checklist...</th>
<th>Strongly Disagree</th>
<th>Neutral</th>
<th>Strongly Agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. is applicable to organizations in different industries.</td>
<td>1 2 3 4 5 6 7 8 9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. is applicable to organizations regardless of structure.</td>
<td>1 2 3 4 5 6 7 8 9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. is applicable to profit and nonprofit organizations.</td>
<td>1 2 3 4 5 6 7 8 9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. takes appropriate factors into account.</td>
<td>1 2 3 4 5 6 7 8 9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. is relevant, fit for evaluating organizational effectiveness.</td>
<td>1 2 3 4 5 6 7 8 9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. is impartial and unbiased.</td>
<td>1 2 3 4 5 6 7 8 9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. does not inappropriately discriminate against organizations.</td>
<td>1 2 3 4 5 6 7 8 9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8. costs are reasonable in terms of time.</td>
<td>1 2 3 4 5 6 7 8 9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9. costs are reasonable in terms of specialist expertise required.</td>
<td>1 2 3 4 5 6 7 8 9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10. costs are reasonable for those submitting the information.</td>
<td>1 2 3 4 5 6 7 8 9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11. is minimally disruptive to the organization being evaluated.</td>
<td>1 2 3 4 5 6 7 8 9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12. is easy to apply.</td>
<td>1 2 3 4 5 6 7 8 9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>13. is easy to understand.</td>
<td>1 2 3 4 5 6 7 8 9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>14. is externally credible, believable to multiple constituencies.</td>
<td>1 2 3 4 5 6 7 8 9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>15. criteria of merit are complete.</td>
<td>1 2 3 4 5 6 7 8 9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>16. criteria of merit are non-overlapping.</td>
<td>1 2 3 4 5 6 7 8 9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>17. criteria of merit are concise.</td>
<td>1 2 3 4 5 6 7 8 9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>18. criteria of merit are confirmable (i.e., measurable or reliably inferable).</td>
<td>1 2 3 4 5 6 7 8 9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>19. criteria of merit can be measured by a common standard.</td>
<td>1 2 3 4 5 6 7 8 9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20. provides clear direction for evaluating organizational effectiveness.</td>
<td>1 2 3 4 5 6 7 8 9</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Please complete the narrative critique on the following page ➔
Please summarize the primary reasons for your ratings by answering the following questions. Your response should highlight the strengths and weaknesses of the Organizational Effectiveness Evaluation Checklist (OEC). If you need additional space, please continue on the back of this page.

**What are the key strengths of the checklist?**

**What are the key weaknesses of the checklist?**

**What is missing from the checklist?**

**What would you suggest or recommend improving?**

We thank you for your participation and greatly appreciate your support for our research.
Appendix C

Sociodemographic Questionnaire
Sociodemographic Questionnaire

Please select one response per question or write your answer in the space provided. All responses will be kept confidential and reported only in the aggregate.

1. What is your gender?
   - Female
   - Male

2. What is your current age?
   - Less than 25
   - 25 – 39
   - 40 – 54
   - 55 – 69
   - 70 or older

3. What is the highest academic degree you have attained?
   - Doctoral degree
   - Masters degree
   - Bachelors degree
   - Other (please indicate):

4. In general, which of the following best characterizes the extent of your professional research experience, if any?
   - None
   - Very little
   - Some
   - A great deal
   - Extensive

5. In general, which of the following best characterizes the extent of your professional evaluation experience, if any?
   - None
   - Very little
   - Some
   - A great deal
   - Extensive

6. In general, which of the following best characterizes the extent of your professional management experience, if any?
   - None
   - Very little
   - Some
   - A great deal
   - Extensive

7. What is your professional or disciplinary area(s) of expertise?
Appendix D

Consent Document
Organizational Effectiveness Evaluation Checklist Expert Panel Review
Consent Document

Principal Investigator: Kenneth C. O'Shaughnessy, Ph.D.
Student Investigator: Wes Martz

You have been invited to participate in a research project titled "Organizational Effectiveness Evaluation Checklist Expert Panel Review." This research is intended to capture critical feedback from an expert panel regarding the Organizational Effectiveness Evaluation Checklist (OEC) in an effort to improve the checklist. This project is part of Wes Martz's dissertation project.

You will be asked to critically review the OEC and provide written comments and suggestions to Wes Martz. In addition, you will be asked to complete a brief critical feedback survey after reviewing the OEC. In all, you will be asked for approximately two hours of your time. Prior to reviewing the checklist, you will be asked to complete a brief socio-demographic questionnaire. The socio-demographic questionnaire will ask you to provide general information about yourself, such as your gender, age, level of education, and research, evaluation, and management experience.

There are no known risks or discomforts associated with this study. However, there is an inconvenience related to the time required to participate.

One way in which you may benefit from this activity is from the knowledge that your participation will make a contribution to the advancement of evaluating organizational effectiveness. In addition, all participants will be offered a copy of the revised OEC and a PowerPoint presentation summarizing the findings. In the longer term, it is hoped that the research will augment existing models for a better understanding of the construct organizational effectiveness, and improve the way in which organizations are evaluated to benefit organizational members and society.

All of the information collected from you is confidential. That means that your name will not appear on any papers on which this information is recorded. All forms will be coded, and Wes Martz will keep a separate master list with the names of participants and the corresponding code.

B-402 Duerink Hall, Kalamazoo, MI 49008-5217
phone: (269) 387-5095, fax: (269) 387-1623

204
numbers. Once the data are collected and analyzed, the master list will be destroyed. All other
forms will be retained for three years in a locked file in the principal investigator's office.

You are free to decide not to participate in this study or to withdraw at any time during the study
without prejudice or penalty. If you have any questions or concerns about this study, you may
contact either Wes Marez at 269-352-4912 or Dr. O'Shaughnessy at 616-742-5028. You may also
contact the chair of Human Subjects Institutional Review Board at 269-387-8293 or the vice
president for research at 269-387-8298 with any concerns that you have.

This consent document has been approved for use for one year by the Human Subjects
Institutional Review Board (HSIRB) as indicated by the stamped date and signature of the board
chair in the upper right corner. Do not participate in this study if the stamped date is older than
one year.

You will receive two copies of the consent document. One is for your records, the other is to
sign and return to Wes Marez.

Your signature below indicates that you have read and/or had explained to you the purpose and
requirements of the study and that you agree to participate.

__________________________  ______________________
Signature                                      Date

Consent obtained by: ____________________________  ______________________
Initials of researcher                          Date
Appendix E

Approval Letter From the Human Subjects Institutional Review Board
Date: February 21, 2008

To: Kenneth C. O'Shaughnessy, Principal Investigator
    Wes Marz, Student Investigator for dissertation

From: Amy Naugle, Ph.D., Chair

Re: HSIRB Project Number: 08-02-23

This letter will serve as confirmation that your research project entitled “Organizational Effectiveness Evaluation Checklist Expert Panel Review” has been approved under the expedited category of review by the Human Subjects Institutional Review Board. The conditions and duration of this approval are specified in the Policies of Western Michigan University. You may now begin to implement the research as described in the application.

Please note that you may only conduct this research exactly in the form it was approved. You must seek specific board approval for any changes in this project. You must also seek reapproval if the project extends beyond the termination date noted below. In addition if there are any unanticipated adverse reactions or unanticipated events associated with the conduct of this research, you should immediately suspend the project and contact the Chair of the HSIRB for consultation.

The Board wishes you success in the pursuit of your research goals.

Approval Termination: February 21, 2009
Appendix F

Permission to Use Copyright Information
March 6, 2008

John Wiley & Sons, Inc.
Permissions Department
111 River Street
Hoboken, NJ 07030-5774

I would like to request your permission to include an excerpt from the following item in my dissertation:


I have created a figure that shows the four quadrants of the competing values framework and the criteria associated with each value orientation shown on the attached page. The information I used in the figure is adapted from page 48, Figure 5 of the text. The source will receive full credit in the manuscript.

For your convenience, I am including a space below for your signature to indicate your permission for my use of the above-mentioned material. By signing below, you give ProQuest Information and Learning (formerly University Microfilms) the right to supply copies of this material on demand as part of my doctoral dissertation.

Please attach any other terms and conditions for the proposed use of this item below. If you no longer hold the copyright to this work, please indicate to whom I should direct my request.

[Signature]

Please return the signed copy of this letter in the postage-paid envelope provided.

Sincerely,

Wes Martz
P.O. Box 190
Portage, MI 49081
Email: wes.martz@wmich.edu

No rights are granted to use content that appears in the Work with credit to another source.