

The Effects of Alternative Reports of Human Resource Development Results on Managerial Support

Brent W. Mattson

Managerial responses to human resource development (HRD) results evaluation reports were experimentally investigated as a function of (1) how evaluation information was presented and (2) reported HRD program impact levels. Managers (n = 233) read a business scenario in which they were asked to make a decision about whether to implement a development program. They were then exposed to one of nine experimental treatment conditions (evaluation report type × reported program impact level). The report types included utility analysis, critical outcome technique, and anecdotal evaluation reports. Results were varied at three impact levels (low, average, and high). Findings of the study showed that managers perceived utility analysis and critical outcome technique reports as almost equally useful in decision making; however, the anecdotal evaluation report was found to be significantly less useful than either of the other two report types. There was no effect of the reported program impact level on the perceived usefulness of the evaluation reports for decision making. Furthermore, there was no interaction between report type and impact level on the perceived usefulness of the reports for decision making. These findings show that managers prefer information about the financial results of HRD interventions to anecdotal information, regardless of the reported level of impact.

Scholars and practitioners of HRD have focused attention on the evaluation of HRD interventions for several decades. This issue has recently been the focus of considerable discussion and debate, fueled by an increasingly competitive,

Note: This article is based on my doctoral dissertation completed at the University of Minnesota. I thank my committee members: Richard Swanson, Gary McLean, Paul Sackett, Ross Azevedo, and Gary Leske. I also thank the General Electric Company. Requests for reprints should be sent to Brent W. Mattson, 9034 Neill Lake Road, Eden Prairie, MN 55347.

global business environment that has pressed HRD practitioners to demonstrate that their interventions contribute directly to the corporate bottom line.

Although several studies have demonstrated the effectiveness of various evaluation techniques in estimating the results of HRD interventions, I found little published research that examines managerial responses to such reports. A few studies show that management responds most favorably to evaluation reports that emphasize business results as opposed to participant reactions, learning, or behavior change. For example, Kusy (1988) surveyed non-HRD managers regarding their references for four different types of HRD evaluation information. He found that subjects showed significantly greater support for results-level evaluation information than for any other type of evaluation (that is, reaction, learning, or behavioral evaluations).

Although results-level program evaluation models differ in method and complexity, they share the assumption that managers emphasize and utilize such information in their operational decisions. For instance, Cascio (1989) asserted that "decision rules are necessary for assessing how sensible it is to invest the firm's dollars in personnel programs, as opposed to capital improvements, financial assets, or other investments" (p. 67). Similarly, Swanson and Holton (1999) noted that training managers believe that the financial evaluation of interventions is important for facilitating managerial decisions.

Over the past several years some investigators have questioned the assumption of managerial use of analytical information in their decision-making processes. These lines of research have implications for the ways in which managers respond to alternative HRD intervention evaluation reports. For instance, seminal research on managerial decision making has generated insight into how managers actually use information. Mintzberg (1973, 1975), in his work on what managers actually do, found that they do not use rational, systematic, or orderly models for decision making; instead, they rely heavily on internal subjective criteria. Mintzberg's work underscores the need to better understand how managers respond to various types of HRD evaluations.

The purpose of this study was to compare three different ways of presenting results-based information about an HRD program to decision makers. The goal was to increase understanding about managerial perceptions to enhance the effectiveness of HRD evaluation practices in organizations.

Decision-Making Framework

As stated above, assessments are tools that HRD practitioners can use to influence managerial decisions regarding HRD interventions. However, there is very little information in the literature regarding managerial responses to HRD assessments. Because HRD assessments, in general, represent a type of analytical information, and because a good deal of literature exists on the subject of how managers and other individuals process such information, it is possible to make use of these insights in terms of background information and

hypothesis formulation (Latham and Whyte, 1994). A framework for addressing the notion of managerial responses to HRD assessment information was synthesized from research in a number of fields of study, including cognitive psychology (Fishbein and Ajzen, 1975), communication and persuasion (Petty and Cacioppo, 1986), marketing (Gotlieb, Schlacter, and St. Louis, 1992), management information systems (Davis, 1989), and human resources (Carson, Becker, and Henderson, 1998).

Fishbein and Ajzen's (1975) theory of reasoned action articulates the definitions of and relationships between beliefs, attitudes, and behaviors, and also how external stimuli are linked to them. According to Fishbein and Ajzen, external stimuli influence an individual's attitude toward a behavior indirectly by influencing his or her salient beliefs about the consequences of performing the behavior. According to Fishbein and Ajzen, an individual's overall attitudinal response toward an external stimulus (for example, a new piece of information, a new tool, or new product) is hypothesized to be a major determinant of whether he or she actually uses the information to make decisions. Attitudinal response toward using information is, in turn, a function of two beliefs: perceived ease of use of the information and perceived usefulness of the information. Perceived usefulness is defined as a quality that causes information to be relevant, important, or valuable to a decision maker. Perceived ease of use is defined as the extent to which information is understandable and unambiguous or clear (Petty and Cacioppo, 1986).

Perceived Usefulness

To the extent that the information contained in an assessment report is seen as valuable or useful, a decision maker may choose whether or not to base a decision on the information. Two factors, cited in a broad range of literature, have been identified as able to influence a person's perceptions regarding the value of information: credibility of the information source (Swanson and Mattson, 1997; Mintzberg, 1975; O'Reilly and Roberts, 1976) and organizational values (Fishbein and Ajzen, 1975; Schein, 1992).

Credibility of Information Source. Credibility is believed to comprise two underlying dimensions: expertise and trustworthiness (Dholakia and Sternthal, 1977). Research on the credibility of information indicates that it may be the trustworthiness of the information source, more than the perceived expertise of the source, that determines whether information is believed. What the substantial body of research on credibility, or source credibility as it is sometimes called, indicates is that a piece of information may be accepted as valid or not valid depending on the degree of trust between the information receiver and sender. In one of the few studies in the HR assessment literature addressing credibility, researchers found that managers were more likely to adopt a selection system when the analytical report describing the potential gains of the system was accompanied by a recommendation from a "trusted adviser" as

opposed to an unknown psychologist's written, videotaped, or personal recommendation (Whyte and Latham, 1997).

Credibility of Reported Impact Levels. Another aspect of credibility that has received limited attention in the literature is that of the credibility of the reported intervention impact level of HRD interventions. Some evidence shows that reported impact levels influence managerial perceptions of the credibility of information. In their investigation of the results of a sales communication training program, Swanson and Mattson (1997) noted that managers were initially skeptical of an 800 percent return on investment (ROI) figure. They noted that the managers were disabused of their skepticism only after a table of outcome values was presented and it was communicated that the organization's comptroller had assisted in the impact-level calculation. Similarly, Carson, Becker, and Henderson (1998) posited that the size of a utility estimate may induce a negative reaction from managers.

Organizational Culture. Another manifestation of the tendency of decision makers to use heuristics is observable in what Schein (1992) refers to as organizational culture. By definition, members of any organizational culture will be subjected, to a greater or lesser degree, to the norms, values, and beliefs of the overall group.

McCall and Kaplan (1990) posited that organizational values are, in and of themselves, a source of information to be interpreted by a manager in any decision-making scenario. They noted that an organization's values directly affect what managers focus on when making decisions. Information relevant to a given value is more likely to reach the manager in more detail and have more implications. Research on cultural factors as they relate to individual behavior indicates a strong linkage between the values and past experiences that organizational members share and what behaviors are accepted or compatible within the culture (Rogers and Shoemaker, 1971; Schein, 1992).

Perceived Ease of Use

Perceived ease of use was defined as the extent to which the assessment information is unambiguous or clear. As with perceived usefulness, a number of studies across a range of disciplines deal with constructs related to perceived ease of use. A good deal of research exists regarding how managers process information. For example, Carson, Becker, and Henderson (1998) noted that there is a strong rationale for the notion that the manner in which information is presented to managers will make a difference in their interpretation and response to it. O'Reilly also noted that even though decision makers recognized information sources as being of high quality, they often used sources that provided lower-quality information but were more accessible. He explained these findings in terms of the costs involved in obtaining information from accessible sources. Because of the environment in which decision makers typically operate, it may have been that they were simply unable to seek out higher-quality information when it came from less accessible sources.

One factor that affects the clarity of information is the degree to which information is abstract or concrete. Assessment reports—particularly those that assess intervention results—rely on the analysis of data. According to O'Reilly (1983), users of such information complain that the reports are too dry, rely too much on statistical analysis, or don't get at the real issues. He posited that such complaints might be indicative of the inability of decision makers to use abstract information. He also noted that users of information often focus on single, concrete, often dramatic examples when making decisions, even when such examples may not be representative of the larger picture.

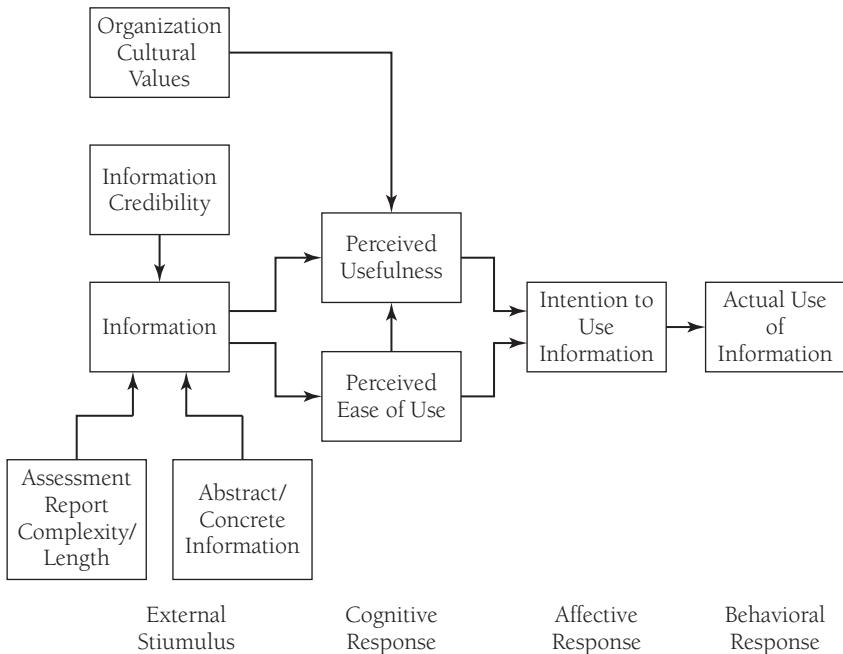
Another factor that is known to affect managerial decision making is information complexity. In their general framework for understanding persuasive communications, Petty and Cacioppo (1986) posited that increasing the complexity and decreasing the comprehensibility of information will increase the likelihood that the information will be negatively evaluated by recipients.

In summary, research on managerial decision making suggests that perceived usefulness and perceived ease of use are fundamental constructs that are important determinants of an individual's attitudes about an external stimulus (that is, an assessment report). Furthermore, a review of the constructs related to managerial use of analytical information indicates several constructs that are relevant to understanding how managers process information: information credibility (trustworthiness, expertise, and estimate size), information complexity (analytic information and report length), whether information is presented in an abstract versus a concrete form (equations, opinions, anecdotal evidence, or business results), and cultural values. Of these four constructs, the first three relate directly to the information source. The fourth, cultural values, is an artifact of the organizational culture and exerts influence on cognitive responses directly. A depiction of these constructs is presented in Figure 1.

Effects on Research Methodology

The research on managerial use of rational decision-making strategies (Mintzberg, 1975; O'Reilly, 1983) has begun to manifest itself in studies regarding managerial use of HR intervention information. Such studies have yielded mixed findings with respect to the use of analytical tools to support managerial decision making. With respect to managerial use of analytical information in an HR context, the majority of research has been conducted on the use of utility analysis for proposed selection systems. However, the findings of the various investigations have implications beyond selection system proposals, and beyond utility analysis. The following sections review the literature of how managerial decision making is related to the three HRD evaluation methods (utility analysis, critical outcome technique, and the anecdotal approach). It is important to note that the three methods do not produce equivalent information. The purpose of the study was not to compare three reports of the financial impact of an intervention but to compare three methods of reporting results-based information that differ in perceived complexity. For example, the

Figure 1. Theory of Reasoned Action Model and Factors That Influence Decision Making



utility analysis method relies exclusively on presenting financial information; the critical outcome technique reports both financial results and nonfinancial results; and the anecdotal approach reports perceptions of results-based information. Furthermore, the fact that the methods behind the creation of each report differ (in terms of ease of preparation) is a useful point of discussion but not the central thrust of the investigation.

Utility Analysis Evaluation Method. According to Cascio (1989), utility analysis is a tool for expressing the outcomes of a training program in terms of expected dollar returns. Utility analysis uses a capital budgeting framework to specify the cost of a given HRD program, the incremental benefits derived during each period, the duration of those benefits, and the discount rate that represents the organization's minimum expected ROI.

Although only a small number of studies have described cases in which utility analysis was used to estimate the impact of an HRD program or programs, the findings have been generally positive. For instance, Mathieu and Leonard (1987) applied utility analysis to an HRD program in supervisory skills. They noted that even with conservative valuations applied to future utility estimates, the benefits of training the employees in this study far outweighed

the costs. They also noted that practitioners and researchers may be skeptical about the results figures.

In recent years however, a good deal of skepticism about the usefulness of the utility analysis technique in influencing managerial decision making has been noted (Carson, Becker, and Henderson, 1998; Hazer and Highhouse, 1997; Latham and Whyte, 1994; Macan and Highhouse, 1994; Whyte and Latham, 1997). Although most of these studies describe managerial reactions to utility analysis with respect to selection programs, their findings are, given the nature of utility analysis, potentially generalizable to the context of HRD evaluation. For example, Latham and Whyte discovered that managers who received utility analysis information regarding a proposed selection practice were less likely to support the practice than if they had received the same proposal without utility analysis information. In their study, they presented 143 members of an executive training program with information that supported a selection practice using various effectiveness criteria. They varied the information in four ways: validity information only, a combination of validity and expectancy table information, a combination of validity and utility information, and a combination of validity, expectancy table information, and utility information. After reading the proposal, managers in each of the four presentation groups responded to an eight-item scale assessing their likelihood of accepting the proposal and of committing organizational resources to its implementation.

With regard to the factors that affect managerial decision making, utility analysis does not fare well. Although it is capable of accounting for many of the factors that affect dollar utility, managers generally find it complex and somewhat abstract. Furthermore, as several studies have found, utility estimates tend to be discounted (Mathieu and Leonard, 1987).

Critical Outcome Technique Evaluation Method. The critical outcome technique (COT) is an evaluation method that can be applied to any performance-improvement or HRD intervention (Mattson, Quartana, and Swanson, 1998a, 1998b; Mattson, 2000; Swanson and Mattson, 1997). Although most scholars and practitioners agree that evaluation should be an integral part of the intervention analysis, design, development, and delivery, in practice this is often not the case (Swanson, 1996). The COT was developed as a means of evaluating intervention effectiveness in a systematic, post hoc manner.

Unlike methods of evaluation that propose categories or taxonomies of potential results of HRD interventions (for example, the four-level model), the COT is a process model for evaluating the outcomes of such interventions. That is, it focuses less on the myriad ways of examining what might have happened as a result of an intervention and more on the business and financial results or outcomes. Critical outcomes can be thought of as business results at the organizational, process, or individual levels or financial results or benefits in terms of money or monetary ratios. Business results that are monetized become financial results.

Several case studies have been published using the COT applied to HRD programs (Mattson, Quartana, and Swanson, 1998a, 1998b; Swanson and Mattson, 1997). However, unlike utility analysis, no published study examines the use of the COT with regard to managerial decision making. Although the COT is capable of evaluating concrete business results in a relatively straightforward manner, some evidence suggests that decision makers tend to discount intervention values (Swanson and Mattson, 1997).

Anecdotal Methods of Evaluation. The literature on decision making as it relates to selection system interventions clearly indicates that decisions are often based on anecdotal information (for example, someone's opinion). As noted above, Mintzberg (1975) found that one common approach managers use when making decisions about allocating resources to a project is to "pick the man rather than the proposal" (p. 59) and that furthermore, managers considered a number of subjectively weighted factors when making a decision. As such, the input of analytical information into a given resource allocation decision was weighted against a number of other factors. The evaluation literature is rife with studies of different evaluation methods and their application in organizations. However, there is a dearth of literature regarding anecdotal approaches to evaluation. One possible reason for this may be that anecdotal approaches tend to be viewed as unsystematic, uninterpretable evaluation studies (Rossi and Freeman, 1993; Morrow, Jarret, and Rupinski, 1997).

Because no definition for an anecdotal evaluation approach was found in the HRD literature, this type of evaluation was defined in this study by its characteristics with respect to the other "generally interpretable" and "generally uninterpretable" designs (Cook and Campbell, 1979). An anecdotal report is one that has little or no basis in actual organizational data. In contrast to other types of designs, it utilizes no controlled sampling procedures; employs limited, possibly inconsistent, instrumentation; uses no control groups or pretests; employs limited or no careful consideration of threats to internal validity; and may contain no cost-benefit analysis component. It is important to underscore that anecdotal evaluation reports should not be confused with qualitative methodology. There is a wealth of information regarding the use of qualitative techniques to evaluate HRD programs (for example, action research and organizational learning). However, the purpose of the study was to compare evaluation methods that differ in the complexity of information presented to decision makers.

In summary, processes and techniques for conducting anecdotal evaluation are rarely addressed in research on HRD evaluation. Given the relative lack of systematic results-level HRD evaluations that occur in organizations and the notion that, in general, managers find it difficult to conceive of HRD interventions in terms of their contribution to the bottom line, it is reasonable to conclude that anecdotal types of evaluation not only occur frequently but are also used as the basis for decisions about the future use of an HRD intervention (that is, to modify, eliminate, continue, or expand). In terms of the research on managerial decision making, anecdotal evaluations may rely heavily on the

credibility of the source of information rather than on technical complexity of the evaluation method or concrete information about the results.

Because of the increasing attention recently being paid to HRD results evaluation and the proliferation of evaluation methods, it is important to question the degree to which different evaluations affect managerial decision making. The three approaches reviewed for this study vary widely with respect to the factors that are known to influence decision making.

Furthermore, no published study was found that directly examines the reported impact level of an HRD intervention on managerial decision making. In the area of utility analysis, Schmidt, Hunter, McKenzie, and Muldrow (1979) noted that practitioners tend to be skeptical of utility analysis forecasts. The primary reason for this skepticism is that report readers have difficulty conceiving of performance change due to training in financial terms. In addition, the magnitude of the estimate of the dollar value of an intervention may have some impact on the perceived credibility of the evaluation method, and hence on any subsequent decision making. Swanson and Mattson (1997) noted that some managers were skeptical of the values attributed to a sales training program using the COT. Because reported financial results values are known to influence managers' perceptions of information in general, it is important to examine how different reported impact levels affect decisions regarding HRD interventions.

Statement of the Problem

Because of the important role of management in determining the organizational and financial support of the HRD function, it is critical to consider how managers respond to results-based HRD evaluations. Research on managerial decision making describes several factors that are likely to influence how an individual perceives and makes use of information. However, HRD professionals currently have little understanding of how managers respond to HRD evaluations. This study was designed to determine the extent to which an HRD results evaluation based on three methods (utility analysis, the COT, and an anecdotal evaluation) influences managers in deciding whether to invest in a training program. More specifically, the study posed two primary questions. First, does the *type of evaluation* used to convey the results of an HRD intervention affect managers' perceptions of the usefulness of the information for decision making? And second, are managers' perceptions of the usefulness of the information for decision making affected by the *reported impact level* of the intervention?

In the light of these questions, the study tested the following hypotheses:

HYPOTHESIS 1. *There will be a difference in managers' perception of the usefulness of the information for decision making based on the three results-evaluation techniques.*

HYPOTHESIS 2. *There will be a difference in managers' perception of the usefulness of the information for decision making based on the reported impact levels.*

HYPOTHESIS 3. *There will be a difference in managers' perception of the usefulness of the information for decision making based on an interaction of the reported impact level and evaluation method.*

Method

Sample. The sample for this study was drawn from a population of managers at a large financial services organization. The organization is a division of a global, diversified manufacturing and service organization. Study materials were distributed to 320 middle managers (first-line supervisors were not included) through the organization's internal mail system. The sample consisted of 233 experienced middle managers (those with at least one year of experience managing others). Of the group, 108 were women and 125 were men with an average age of 36.93 years. The study sample size of 233 increased power for the interaction of impact-level and evaluation report types to .87. For the main effects of evaluation report type and impact level, the sample of 233 increased statistical power to .93.

Regarding work experience, 6 individuals had work experience in HR and 56 had experience in the finance function. The remaining 172 managers worked in other functions (for example, sales, operations management, information technology, risk management, and legal services). The managers had an average of 7.04 years of experience managing others. Middle managers were chosen because it was believed that they were likely to have made decisions regarding training programs (for example, whether to send employees to the program and whether to allocate departmental resources in order for their direct reports to attend).

Study Design. A 3×3 (report type \times impact level) randomized experimental design was used to determine the effect of the differences in type of evaluation report and impact level on a manager's intention to use the information in decision making.

Nine sets of stimulus materials were used in this study. All of the materials contained a letter describing the research, instructions for completion, a description of the decision scenario, a report based on one of the three HRD results evaluation methods and one of three impact-level values, and a twenty-five-item instrument. Each subject received only one of the nine versions of the stimulus materials. Subjects were randomly assigned to one of the nine experimental conditions.

The instructions informed subjects that they were participating in a study about HRD decision making. The managers read a vignette requiring them to imagine that they were a senior executive in their business unit. They were informed that other key executives were concerned about the ability of employees in the organization to manage their time effectively. Furthermore, they were

informed that there was an effort underway to determine if a time-management program, implemented in another company business unit, should be implemented in their business unit. They were also provided with information about the projected costs of implementing the program in their organization. Beyond this description, the managers were exposed to one of the nine evaluation-report conditions (that is, training results reports). Table 1 illustrates the various independent variable conditions and the number of subjects in each condition.

Time management training was selected as the program for the scenario because it is a common type of skill-building program. Because the materials were distributed to managers in several different business functions, the program described needed to have broad applicability across all types of employees and be generally viewed as transferable to the job. Time management skills had never been addressed in any uniform way across the organization. As Morrow, Jarret, and Rupinski (1997) noted, time management programs can be considered management development programs and are typically viewed as easily transferable and capable of helping participants to achieve visible results.

Credibility was varied solely by changing the reported impact levels in each evaluation. A “training manager” authored each report in order to minimize variance in perceived usefulness due to the subject’s perceptions of the credibility of the information source.

Instrument Design

The research instruments were composed of the nine results-evaluation reports and the twenty-five-item instrument. Prior to designing the instruments, the researcher explored the topic of results evaluation in a literature review. The processes through which each instrument was designed are described below.

Utility Analysis Report. As noted, Cascio (1989) asserted that utility information should be presented to managers in a concrete manner. However, an extensive review of the literature and personal conversations with several scholars of utility analysis yielded no prescribed format for presenting the findings of a utility study. To minimize experimenter bias, the format of the utility report for this study was based on earlier studies using utility analysis to estimate the value of selection procedures. Subjects were provided

Table 1. Sample Size of Independent Variable Conditions

Report Type	Impact Level			Total
	Low	Average	High	
Anecdotal	30	32	15	77
COT	22	30	28	80
Utility analysis	22	32	22	76
Total	74	94	65	233

with a two-page report containing (1) a summary statement of the program's effectiveness, (2) a basic description of the utility analysis procedures used to derive the financial impact of the program, and (3) a statement summarizing the approximate financial gain of the program in dollars and ROI. Furthermore, on page 2 of the evaluation report, detailed information regarding the calculation of the utility of the program was presented. Specifically, the model was shown and each of its terms in the model described (Hazer and Highhouse, 1997; Latham and Whyte, 1994). The report explained that the estimates for each term in the utility model were obtained from research studies and the organization's personnel records. When these estimates were entered into the utility equation, the resulting dollar gain calculated was described as the approximate net returns from the program over a one-year time period.

COT Report. In the COT condition, I provided a two-page report containing (1) a summary statement of the program's effectiveness, (2) a basic description of the COT procedures used to derive the value of the program, (3) a summary of the results attributed to the program in the other business unit, and (4) a statement summarizing the approximate financial gain in dollars and ROI. Furthermore, on page 2 of the report, detailed information about each of the outcomes attributed to the program by participants in the other business unit (that is, number and types of sales, employee retention figures, and improvement in customer service levels). Each of these outcomes was summarized in a table and in a brief paragraph describing the calculations. This report format is based on the prescribed format for presenting results of a COT analysis to program stakeholders (Mattson, 2000; Swanson and Mattson, 1997).

Anecdotal Evaluation Report. Because I found no published format for describing the results of an anecdotal evaluation, a report was created on the basis of information from three HRD experts. The experts were two professors on the investigator's dissertation committee and one HRD leader in the organization in which the study was conducted. Each expert was asked to describe the key elements one would expect to see in a report that was based on an anecdotal method (for example, no uniform data collection method, using only anecdotal comments as evidence). Expert opinions were remarkably consistent. To ensure consistency, each expert was presented with multiple drafts of the anecdotal report as it was developed and modified. One expert suggested that the report read more like a memo than a rigorous evaluation report to approximate the unsystematic nature of the "evidence" of program results. The final written report was less than one page long and reflected the conversational nature of the anecdotal reporting style. Each expert independently agreed that it was a good representation of an anecdotal report.

Impact-Level Conditions. Impact level was deemed to be an important criterion because impact values were found to influence the perceived credibility of information (Carson, Becker, and Henderson, 1998; Swanson and Mattson, 1997). Specifically, it was believed that impact values that were

significantly higher than what managers might reasonably expect to find in an evaluation report might cause them to discount the information contained in the report. In each of the three evaluation-report conditions, the reported program impact level was presented as one of three values—low, average, or high. In order to determine the appropriate values for each of the three program impact levels, I consulted a small group of nontraining managers ($n = 12$). These managers were asked to estimate the impact of the time management skills training program in quantitative terms. Specifically, each manager received an oral description of the training program objectives from the investigator, an outline of the course content, and a profile of prospective participants. Managers were asked to provide a dollar value estimate of the value of the performance outcomes that might reasonably be expected as a result of the implementation of the time management skills program over a one-year period for a group of 250 sales and service managers. Next, managers were asked to provide an estimate of the highest value that might possibly be expected. Finally, they were asked to provide an estimate of the lowest positive value that might possibly be expected. Estimates of the values of the outcomes ranged from \$20,000 to \$4,000,000 (mean estimate = \$600,000) for the “reasonable” program value. Estimates for the “highest expected value” ranged from \$800,000 to \$10,000,000 (mean estimate = \$4,200,000). Estimates for the “lowest expected value” ranged from \$1 to \$250,000 (mean estimate = \$174,000). Estimates were collected and collated and incorporated into the reports. These estimated values correspond to ROI values of 400 percent, 2,800 percent, and 116 percent for the average, high, and low expected program outcome values, respectively.

In each of the utility analysis and COT reports the terms of the equations used to derive the financial values were varied consistently within each type of report to compute the appropriate ROI figure. The ROI figures were used as the basis of the program outcome values (impact levels) in both the COT and utility analysis evaluation reports, respectively.

Because the anecdotal evaluation report contained no quantitative values, only descriptive statements of the effectiveness of the program, key modifiers were changed to represent low, average, and high program effectiveness ratings. For example, in the “high” condition, the modifiers *extremely*, *very highly*, *very significantly*, and *most* were used consistently to refer to enrollment levels, general program ratings, and the number of participants who reported the opportunity to use the skills learned in training back on the job, respectively. In the “low” condition, these modifiers were replaced with *good*, *fairly well*, and *some*, respectively.

Decision-Making Instrument. After reading the scenario and the evaluation report, the subjects responded to an instrument containing twenty-five items. The first ten items were demographic in nature, the remaining fourteen items tapped subjects’ perceptions of the information contained in the evaluation report, and a final item asked which page of the evaluation report

the subject used to make his or her ratings. The fourteen items that tapped subjects' perceptions of the evaluation report were grouped into two scales, or indices. Items on the scales were adapted from items in two similar studies (Hazer and Highhouse, 1997; Whyte and Latham, 1997) in the literature on the managerial decision making related to selection systems.

The first scale, the perceived usefulness measure, contained twelve items designed to assess the subjects' intention to use the information contained in the evaluation report in decision making. The scale addressed issues such as subjects' commitment to implementing the program, their confidence in the ability of the program to generate results in their organizations, and their perceptions of their ability to justify to others their decision whether to implement the program. The scale also addressed issues such as the trustworthiness of the information in the report, their faith in the information, and their confidence in the results described in the report. The managers responded to the items using a five-point Likert-type scale, where 1 = strong disagreement with the statement and 5 = strong agreement with the statement. The internal reliability coefficient (Cronbach's alpha) for the entire scale was .94.

The second scale contained two items designed to measure subjects' understanding and their perceptions of the clarity of the information presented in the evaluation reports (that is, perceived ease of use). The items on these scales were similar to those used by Whyte and Latham (1997) and Hazer and Highhouse (1997). The internal reliability coefficient (Cronbach's alpha) for the entire scale was .87. Both sets of items were subjected to a common factor analysis using a principal-axis extraction procedure with oblique rotation. These analyses confirmed that one construct existed for each scale.

Results

The cell means and standard deviations for the sum of responses to all twelve perceived usefulness scale items are presented in Table 2 for the 3 × 3 (report type × impact level) design. Perceived usefulness scale scores ranged from 12 through 60, with a higher number indicating a stronger inclination to use the information contained in the report to make a decision to implement the

Table 2. Mean Perceived Usefulness Index Scores and Standard Deviations by Report Type and Impact Level

Report Type	Impact Level			Overall M(SD)
	Low M(SD)	Average M(SD)	High M(SD)	
Anecdotal	3.00 (.81)	2.81 (.64)	2.93 (.63)	2.91 (.71)
COT	3.61 (.45)	3.48 (.56)	3.54 (.46)	3.54 (.49)
Utility analysis	3.43 (.78)	3.60 (.61)	3.20 (.61)	3.43 (.67)

program. Perceived usefulness scale scores are presented as averages (total perceived usefulness index score divided by twelve) throughout this analysis to facilitate interpretation. Overall, these data demonstrate that the perceived usefulness of the evaluation report was strongest in the COT and utility analysis conditions and weakest in the anecdotal condition. Mean scores on the perceived usefulness index were 2.91 in the anecdotal condition; 3.43 in the utility analysis condition, and 3.54 in the COT condition. Mean scores on the perceived usefulness index by impact level ranged from 3.22 in the high impact level condition, to 3.29 in the average impact level condition, and to 3.35 in the low impact condition.

Perceptions of Usefulness of Results Report Information

Hypothesis 1 stated that there would be a difference in managers' perception of the usefulness of the information for decision making based on the three results-evaluation methods. To test this hypothesis, an analysis of variance (ANOVA) was conducted for the 3 × 3 (report type × impact level) design on managers' responses to the perceived usefulness index. The findings yielded support for this hypothesis. Specifically, the evaluation report type had a statistically significant effect on perceived usefulness $F(2,224) = 19.94, p < .00$.

Pairwise comparisons between all possible combinations of report types were conducted to investigate the effect that the type of report had on the managers' perception of the usefulness of the assessment report. Tukey's studentized range (HSD) test is appropriate for these purposes (Anderson, Sweeney, and Williams, 1993). The comparisons, which are summarized in Table 3, indicated no significant difference between the COT and utility analysis reports. In contrast, the differences between the anecdotal report and the COT and utility analysis reports, respectively, were significant. Specifically, the level of agreement with items on the perceived usefulness scale was lower when the information was reported anecdotally.

The standardized effect sizes calculated for each comparison are also shown in Table 3. The effect sizes of the statistically significant comparisons between the anecdotal report and both the COT and utility analysis report conditions, respectively, were .94 and .75. According to convention, the

Table 3. Pairwise Comparisons of Perceived Usefulness Index Scores

<i>Comparison</i>	<i>Difference</i>	<i>Effect Size Index</i>	<i>Significance</i>
A – B	.10	.01	n.s.
A – C	.63	.94	.05
B – C	.53	.75	.05

Note: n.s.: nonsignificant. A: COT report. B: utility analysis report. C: anecdotal report.

operational definition of small, medium, and large differences between means corresponds to standardized effect sizes of .2, .5, and .8, respectively (Cohen, 1977). The effect size associated with the decrease in perceived usefulness of the anecdotal report in decision making is therefore large with respect to the COT report, and medium with respect to the utility analysis report.

To examine the potential effects of moderating variables on scores on the perceived usefulness index, I conducted a series of analyses. The findings of correlation analyses between the perceived usefulness index and age, experience in finance, experience in human resources, and managerial experience are presented in Table 4.

A Pearson product-moment correlation revealed a significant negative relationship between years of experience in a finance role and scores on the perceived usefulness index, $r(233) = -.25, p < .000$. Based on this finding, I considered the possibility that experience in a finance role might have influenced the effects of the independent variables on manager reactions. The findings of a two-way ANCOVA using finance experience as a covariate are presented in Table 5. Table 5 shows that assessment report type remained significant while neither impact level nor the interaction between report type and impact level was significant when the years of experience in a finance role was controlled. Therefore, experience in a finance role does not explain the differences in perceived usefulness index scores.

Table 4. Correlations Between Moderating Variables and the Perceived Usefulness Index

<i>Variable</i>	<i>Perceived Usefulness Index</i>
Age	.04
Finance experience	-.25**
Human resources experience	.01
Managerial experience	.12

** $p < .001$.

Table 5. Source Table for Two-Way ANCOVA: Effects of Financial Experience on Perceived Usefulness Index

<i>Source</i>	<i>Df</i>	<i>MS</i>	<i>F</i>
Financial experience (covariate)	1	5.55	14.94*
Report type (R)	2	7.50	20.19*
Impact level (V)	2	.43	1.16
R × V	4	.63	1.70
Error	223	(.47)	

* $p < .001$.

Hypothesis 1: Effects of Categorical Moderating Variables. To examine the potential effects of other moderating variables on scores on the perceived usefulness index, the investigator conducted a series of ANOVAs. The findings of ANOVAs of the familiarity variables (familiarity with HR accounting, utility analysis, financial analysis of training, analysis of ROI and return on equity (ROE), and statistical analysis), gender, the perceived ease-of-use index, and whether or not the subject completed the instrument via e-mail or as part of an intact group are presented in Table 6.

Based on these findings, I considered the possibility that familiarity with the financial analysis of training or familiarity with HR accounting might have influenced the effects of the independent variables on perceived usefulness. The findings of two two-way ANCOVAs using these two moderators as covariates are presented in Table 7 and Table 8, respectively. Table 7 shows that assessment report type remained significant while neither impact level nor the

Table 6. Source Table for ANOVA by Moderating Variables

<i>Source</i>	<i>Df</i>	<i>MS</i>	<i>F</i>
Familiarity with human resource accounting	4	1.24	2.73*
Familiarity with utility analysis	4	.90	1.95
Familiarity with financial analysis of training	4	1.91	4.29*
Familiarity with analysis of return on investment (ROI) and return on equity (ROE)	4	.48	1.02
Familiarity with statistical analysis	4	.80	1.71
Gender	1	.65	1.38
Perceived ease of use index average	4	3.98	9.74**
Instrument response: e-mail vs. group	1	.56	1.20

* $p < .05$; ** $p < .001$.

Table 7. Source Table for Two-Way ANCOVA: Effects of Familiarity with Financial Analysis of Training on Perceived Usefulness Index

<i>Source</i>	<i>Df</i>	<i>MS</i>	<i>F</i>
Familiarity with financial analysis of training (covariate)	1	7.18	19.72*
Report type (R)	2	8.80	24.16*
Impact level (V)	2	.61	1.68
R × V	4	.61	1.82
Error	223	(.36)	

* $p < .001$.

Table 8. Source Table for Two-Way ANCOVA: Effects of Familiarity with Human Resource Accounting on Perceived Usefulness Index

<i>Source</i>	<i>Df</i>	<i>MS</i>	<i>F</i>
Familiarity with human resource accounting (covariate)	1	3.67	9.65*
Report type (R)	2	7.71	20.29*
Impact level (V)	2	.52	1.34
R × V	4	.69	1.82
Error	223	(.38)	

* $p < .001$.

interaction between report type and impact level was significant when familiarity with the financial analysis of training programs was controlled. Therefore, familiarity with the financial analysis of training programs does not explain the differences in perceived usefulness index scores.

Table 8 shows that assessment report type remained significant while neither impact level nor the interaction between report type and impact level was significant when familiarity with the HR accounting was controlled. Therefore, familiarity with HR accounting does not explain the differences in perceived usefulness index scores.

The ANOVA findings in Table 6 revealed a significant relationship between scores on the perceived ease-of-use index and scores on the perceived usefulness index. Based on this finding, I considered the possibility that subjects' understanding of the information contained in the assessment reports might have explained the effects of the independent variables on perceived usefulness. The findings of a two-way ANCOVA using the perceived ease-of-use index as a covariate are presented in Table 9. Table 9 shows that assessment report type remained significant while neither impact level nor the interaction between report type and impact level was significant when the perceived ease-of-use index was controlled. Therefore, perceived ease of use does not explain the differences in perceived usefulness index scores.

Table 9. Source Table for Two-Way ANCOVA: Effects of Perceived Ease of Use Index on Perceived Usefulness Index

<i>Source</i>	<i>Df</i>	<i>MS</i>	<i>F</i>
Perceived ease of use index (covariate)	1	8.74	24.64*
Report type (R)	2	5.56	15.55*
Impact level (V)	2	.527	1.47
R × V	4	.387	1.08
Error	223	(.357)	

* $p < .001$.

Hypothesis 2. The second hypothesis stated that there would be a difference in managers' perceptions of the usefulness of the information for decision making based on the reported impact levels. The findings of the ANOVA indicated no statistically significant effect due to the impact level $F(2,224) = .63, p < .54$ on managers' perceptions of the usefulness of the report for decision making.

Hypothesis 3. The third hypothesis stated that there will be a difference in managers' perceptions of the usefulness of the information for decision making based on an interaction of the reported impact level and evaluation method. The findings of the ANOVA indicated no support for the rejection of the third hypothesis. Specifically, there was no report type \times impact level interaction $F(2,224) = 1.44, p < .22$.

Summary of Findings

As noted, the first hypothesis was the only one of the three hypotheses not to be rejected. In other words, the type of evaluation report accounted for a significant difference in the perceived usefulness of the information for decision making. Pairwise comparisons revealed a significant difference in terms of managers' perceptions of the usefulness of the report and between the anecdotal report and both the COT and utility analysis reports. Furthermore, analyses of covariance demonstrated that none of the potentially moderating variables had any effect on the relationship between report type and managers' perceptions of the usefulness of the reports for decision making. Findings of statistical analyses yielded no support for either the second or third hypothesis. Either the reported impact level or the interaction between reported impact level and report type did not affect managers' perceptions of the usefulness of the reports for decision making.

Discussion

Evaluation Report Type. In accordance with the first hypothesis, I found that the type of evaluation report managers received did affect their perception of the information contained in the report. Managers perceived the report based on anecdotal information as least useful for decision making. These findings are inconsistent with the theoretical propositions reviewed in the introduction to this chapter (O'Reilly, 1983; Petty and Cacioppo, 1986). In a complex decision environment, decision makers attend to and use information more when it is presented in a manner that is shorter in length and higher in comprehensibility. The simplistic reporting of anecdotal information and short report length might have influenced perceived usefulness by increasing the comprehensibility of the evaluation report. Conversely, the relative length and technical complexity of both the utility analysis and COT reports should have rendered them less easy to understand, and hence less useful for decision

making. However, this was not the case. Managers found both utility analysis and COT evaluation reports significantly more understandable than the report-based anecdotal information. Furthermore, both COT and utility analysis evaluation reports received significantly more support from managers in terms of their perceived usefulness.

Although both the COT and utility analysis reports fared better than the anecdotal report in influencing managerial perceptions, neither was overly influential. Mean perceived usefulness scores were lower than expected, particularly for the COT report. This finding should not be overlooked. For both evaluation report (that is, COT and utility analysis) scenarios, managers perceived the usefulness of the information to be between 3.5 and 3.75 on a five-point scale, the midpoint of which is 3. This indicates that managers were not particularly impressed with the proposed training program based on the results described in either evaluation report. These findings are relatively consistent with Latham and Whyte (1994) or Carson, Becker, and Henderson's (1998) findings with respect to utility analysis and decision making in the context of HR selection system proposals. Both of these studies employed similar instrumentation and found mean results in the range of 3 to 3.3 on a five-point scale in terms of perceived usefulness.

Perceived Ease of Use. It was anticipated at the outset of this investigation that the COT evaluation report would prove more useful to managers in decision making due to the relative concreteness of the information contained in the report. The utility analysis report contained an algebraic equation and explanation of each of the terms, whereas the COT report described actual business results that were ascribed to the training program. Furthermore, given the findings of studies that examined utility analysis in the context of HR selection system proposals (Latham and Whyte, 1994), I anticipated that managers would perceive the COT report as significantly more useful in decision making than the utility analysis report.

One possible explanation for this finding is that presenting utility information to managers in a user-friendlier manner results in some improvement in the perceived usefulness of the information. Furthermore, managers were not put off by the size of the utility estimate in the high impact level condition because there is no significant difference in perceived usefulness between any of the impact levels.

Organizational Culture

It is possible that the culture of the organization in which this study was conducted could have affected these findings. For example, the organization is known for its extensive use of statistical reporting techniques associated with its Six Sigma quality approach. Managers in the organization are expected to participate in Six Sigma training and use quality tools to measure many aspects of their (and their team's) performance. This may be one reason that none of

the evaluation techniques fared overwhelmingly well in terms of influencing managers' perceived usefulness of the reports. Perhaps none of the reports satisfied managers' shared assumptions of what information and analyses an HRD evaluation report should contain. This may explain the significant difference between the anecdotal evaluation report and both the COT and utility analysis reports. Both the COT and utility analysis reports emphasized the collection and analysis of data to explain the results of the training program. Data collection and analysis are integral to many managers' operating processes. Indeed, the company has been described as having a data-driven culture.

Another well-known artifact of the organization's culture is that the organization places a high value on learning. Because of this, it is conceivable that managers might support the implementation of an HRD program simply because it fosters learning, irrespective of the results that may be expected. However, if this were the case, the anecdotal evaluation report probably would have been received more favorably (that is, would have had a higher perceived usefulness).

Impact Level

The reported impact level attributed to the program did not significantly affect managerial perceptions of the usefulness of the information. This was surprising in light of the findings of Swanson and Mattson (1997), who reported encountering resistance when they presented stakeholders with a report of an 800 percent ROI for a sales communication training program. It is also inconsistent with the proposition of Schmidt, Hunter, McKenzie, and Muldrow (1979) that managers tend to be skeptical of large utility estimates. This is more surprising when one considers that the financial results in the high impact level condition in this study were reported to be equivalent to a 2,800 percent ROI. Furthermore, the financial results reported in the average impact level condition in this study were reported to be equivalent to a 400 percent ROI. This finding should be welcome news to individuals who are charged with assessing the results of training programs. In effect, the study shows that managers were not put off by the size of the impact at any level (that is, low, average, or high).

Conclusions

A 1996 report by the American Society for Training and Development (ASTD) of HRD practice reported that only 3 percent of development programs were assessed in terms of financial impact (Bassi, Benson, and Cheney, 1996). This is surprising given the attention devoted in the literature to the importance of evaluating the results of HRD interventions. This study was designed based on a need for understanding of how managers use results-based information to make decisions about HRD programs. Although similar work has been done

in the area of HR selection, this study provides an understanding of this phenomenon for HRD professionals.

In the study, three results-based reports were used to communicate the results of an HRD intervention. It is clear from its findings that HRD practitioners and consultants should be encouraged to look for ways to include effectiveness data in their proposals for interventions. Furthermore, practitioners must consider which type of evaluation makes the most sense for their organizations based on the information needs of their decision makers. Although utility analysis is a viable means of evaluating the potential effectiveness of an HRD intervention, it is relatively difficult to use and involves estimating a number of factors to derive the final dollar-value utility of an intervention. In contrast, the COT is designed for use after an intervention has occurred. As noted, it is relatively simple to use and can report nonfinancial results (for example, business results) as well as financial results. Although this could prove a drawback in terms of demonstrating that a program caused a change in results, in most cases it should be adequate for producing good enough information for decision makers. Practitioners who use either of these techniques must weigh the costs and benefits of each technique according to their organization's needs and their ability and resources to conduct either type of evaluation. Practitioners would also be wise to work to build acceptance of their evaluation methods with stakeholders prior to providing effectiveness data. By providing basic information to key decision makers prior to reporting results, practitioners could build understanding of the nature of results evaluations. If decision makers are more comfortable with the notion that reports of program effectiveness are similar to other financial estimates, in that all estimates contain errors, they may be more willing to view such reports as useful for decision making.

The tradeoffs described above represent a consideration of pragmatic issues related to HRD evaluation (that is, which technique is best given quality, time, and cost considerations). There is another side of this issue that is rarely discussed: the ethical considerations involved in selecting one approach over another. For example, is it ethical to create a report in a way that does not ensure that a program really contributed to the results? For these types of questions there can be no easy answers. Clearly, a technique that controls for threats to internal validity in a systematic way will garner "better" information than one that does not. However, such precautions are not foolproof. For instance, Sackett and Mullen (1993) assert that "a pre-experimental [generally uninterpretable] design, paired with careful investigation into the plausibility of various threats, is still better than no evaluation at all, given that organizations must make decisions about future training efforts with or without evaluation data" (p. 621). Perhaps the real ethical issue is not whether an experimental evaluation design is preferable to a preexperimental one, but whether or not one chooses to base decisions about HRD programs on systematically gathered information.

The findings of this study demonstrate that managers do indeed find results influential in their decision-making processes. If the findings of the ASTD study (Bassi, Benson, and Cheney, 1996) are any indication, HRD professionals have a long way to go in terms of changing the expectations of the stakeholders of their interventions. When organizations begin to demand greater accountability on the part of their HRD providers not only will the process of providing HRD interventions improve but so, too, may the results.

Future research should examine ways in which proposals for HR interventions can be presented in an optimal fashion. As noted, the mean perceived usefulness ratings leave room for improvement. A qualitative data-gathering technique might be useful to help determine what managers actually think about when they review proposals for HR interventions.

Future investigators may also seek to understand if the information desired by decision makers about HRD interventions varies by organization or function. For example, are decision makers in sales roles persuaded by nonsales HRD intervention results? It is highly likely that decision makers in different organizations (and hence different organizational cultures) would demonstrate different preferences for certain types of information. Therefore, it would be reasonable to study evaluation information preferences across multiple organizations.

Future research may also be appropriate to address the limitations of this study. For example, examining managers in only one company limits generalizability of the findings to that company only. Furthermore, although real managers were used as subjects, the decision-making scenario was artificial. Managers in the study did not have incentives (financial or otherwise) for making the appropriate decisions. Conducting a similar study under real decision-making conditions may change the findings.

Another limitation of the study lies in the presentation of anecdotal information to nontraining managers. As noted, an anecdotal report is one that relies on no real methodology, quantitative or qualitative, to provide data on which a manager can base a decision. Although we know that decisions about HRD programs are frequently made on the basis of gut feel or a hallway conversation with someone who said positive things about how a program helped him or her do a better job, the anecdotal report may not have been an accurate representation of this phenomenon. Nevertheless, the finding cannot be ignored that information about results based on a reasonable data collection strategy was significantly more effective in this case.

Finally, it might be useful to conduct similar studies across different cultural groups to test for differences in the way in which different cultures respond to the presentation of different types of results information.

In conclusion, this chapter addresses an important, but overlooked, area of program evaluation. There are many other evaluation techniques and other methods that can be used to test their effectiveness and the impact of those techniques on managerial decisions. Regardless of the methodologies employed

by future studies, resolution of the important questions surrounding the practice of HRD evaluation will be best achieved through continued empirical exploration rather than conceptual speculation.

References

- Anderson, D. R., Sweeney, D. J., & Williams, T. A. (1993). *Statistics for business and economics*. St. Paul: West.
- Bassi, L., Benson, G., & Cheney, S. (1996). *Trends: Position yourself for the future*. Alexandria, VA: ASTD Press.
- Carson, K. P., Becker, J. S., & Henderson, J. A. (1998). Is utility really futile? A failure to replicate and an extension. *Journal of Applied Psychology*, 83, 84–96.
- Cascio, W. F. (1989). Using utility analysis to assess training outcomes. In I. L. Goldstein (Ed.), *Training and development in organizations*. San Francisco: Jossey-Bass.
- Cohen, J. (1977). *Statistical power analysis for the behavioral sciences*. Orlando, FL: Academic Press.
- Cook, T. D., & Campbell, D. T. (1979). *Quasi-experimentation: Design and analysis issues for field settings*. Skokie, IL: Rand McNally.
- Davis, F. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, 12 (4), 319–340.
- Dholakia, R. R., & Sternthal, B. (1977). Highly credible sources: Persuasive facilitator or persuasive liabilities? *Journal of Marketing*, 13, 416–427.
- Fishbein, M., & Ajzen, I. (1975). *Belief, attitude, intention, and behavior: An introduction to theory and research*. Reading, MA: Addison-Wesley.
- Gotlieb, J. B., Schlacter, J. L., & St. Louis, R. D. (1992). Consumer decision making: A model of the effects of involvement, source credibility, and location on the size of the price difference required to induce consumers to change suppliers. *Psychology and Marketing*, 9 (3), 191–208.
- Hazer, J. T., & Highhouse, S. (1997). Factors influencing managers' reactions to utility analysis: Effects of SDy method, information frame, and focal intervention. *Journal of Applied Psychology*, 82 (1), 104–112.
- Kusy, M. (1988). The effects of types of training evaluation on support of training among corporate managers. *Performance Improvement Quarterly*, 1 (2), 23–30.
- Latham, G. P., & Whyte, G. (1994). The utility of utility analysis. *Personnel Psychology*, 4, 31–46.
- Macan, T. H., & Highhouse, S. (1994). Communicating the utility of human resource activities: A survey of I/O and HR professionals. *Journal of Business and Psychology*, 8, 425–436.
- Mathieu, J. E., & Leonard, R. L. (1987). Applying utility concepts to a training program in supervisory skills: A time-based approach. *Academy of Management Journal*, 30 (2), 316–335.
- Mattson, B. W. (2000). Development and validation of the critical outcome technique. *Human Resource Development International*, 3 (4), 465–488.
- Mattson, B. W., Quartana, L. J., & Swanson, R. A. (1998a). Assessing the business results of management development using the critical outcome technique at CIGNA corporation. In R. Torraco (Ed.), *Academy of Human Resource Development 1998 Annual Proceedings* (pp. 44–50). Atlanta, GA: Academy of Human Resource Development.
- Mattson, B. W., Quartana, L. J., & Swanson, R. A. (1998b). Assessing the business results of management development using the critical outcome technique. In J. J. Phillips (Ed.), *In action: Measuring ROI*. Alexandria, VA: ASTD.
- McCall, M. W., & Kaplan, R. E. (1990). *Whatever it takes: The realities of managerial decision making*. Upper Saddle River, NJ: Prentice Hall.
- Mintzberg, H. (1973). *The nature of managerial work*. New York: HarperCollins.
- Mintzberg, H. (1975, Jul.-Aug.). The manager's job: Folklore and fact. *Harvard Business Review*, 49–61.
- Morrow, C. C., Jarret, M. Q., & Rupinski, M. T. (1997). An investigation of the effect and economic utility of corporate-wide training. *Personnel Psychology*, 50, 91–119.

- O'Reilly, C. A. (1983). The use of information in organizational decision making: A model and some propositions. *Research in Organizational Behavior*, 5, 103–139.
- O'Reilly, C. A., & Roberts, K. (1976). Relationship among components of credibility and communication behaviors in work units. *Journal of Applied Psychology*, 61, 99–102.
- Petty, R. E., & Cacioppo, J. T. (1986). *Communication and persuasion: Central and peripheral routes to attitude change*. New York: Springer-Verlag.
- Rogers, E. M., & Shoemaker, F. F. (1971). *Communication of innovations: A cross cultural approach*. New York: Free Press.
- Rossi, P. H., & Freeman, H. E. (1993). *Evaluation: A systematic approach*. Thousand Oaks, CA: Sage.
- Sackett, P. R., & Mullen, E. J. (1993). Beyond formal experimental design: Towards an expanded view of the training evaluation process. *Personnel Psychology*, 46, 613–627.
- Schein, E. H. (1992). *Organizational culture and leadership*. San Francisco: Jossey-Bass.
- Schmidt, F. L., Hunter, J. E., McKenzie, R. C., & Muldrow, T. W. (1979). Impact of valid selection procedures on workforce productivity. *Journal of Applied Psychology*, 64, 609–620.
- Swanson, R. A. (1996). *Performance-learning-satisfaction evaluation system: The application of the three domain evaluation model to performance improvement, human resource development, organization development, and personnel training and development*. St. Paul: University of Minnesota Human Resource Development Research Center.
- Swanson, R. A., & Holton, E. F. (1999). *Results: How to assess performance, learning, and perceptions in organizations*. San Francisco: Berrett-Kohler.
- Swanson, R. A., & Mattson, B. W. (1997). Development and validation of the critical outcome technique. In R. Torraco (Ed.), *Academy of Human Resource Development 1997 Annual Proceedings*, 64–71.
- Whyte, G., & Latham, G. (1997). The futility of utility analysis revisited: When even an expert fails. *Personnel Psychology*, 50, 601–610.



