

# Kirkpatrick Plus: Evaluation and Continuous Improvement with a Community Focus

□ Ryan Watkins  
Doug Leigh  
Rob Foshay  
Roger Kaufman

*For almost 40 years, Donald Kirkpatrick's framework for evaluation has been used as a basic model for the identification and targeting of training-specific interventions in business, government, military, and industry alike. By approaching evaluation from four different perspectives—reaction, learning, behavior, and results—the model has provided a solid basis for the examination of training's impact on the organization. Despite the current practice of measuring one's success according to the success of one's clients, proposed changes in the model have not been frequently adopted. It is therefore likely time for professionals to reevaluate the utility and responsiveness of the Kirkpatrick framework to meet the value-added requirements of today's organizations. This article identifies tools and concepts for being responsive to the new organizational realities not originally addressed by the Kirkpatrick model.*

□ In 1959, Donald Kirkpatrick (1959a; 1959b; 1960a; 1960b) introduced a framework for the evaluation of training and education which has become a standard in many applications of business, government, military, and industrial training. Kirkpatrick's framework offers performance-improvement professionals, as well as evaluators, a basic model for identifying and targeting training-specific evaluation efforts. The framework urges practitioners to evaluate training or educational programs, or both, at four levels or steps: reaction, learning, behavior, and results (Kirkpatrick, 1994; 1996). Despite changes in the internal and external organizational environments during the past 40 years, revisions to the initial model—such as those proposed by Kaufman, Keller, and Watkins (1995) and Philips (1996)—have not yet been widely adopted. One of the primary changes in organizational environments is increasing demands on individuals and organizations to demonstrate value-added to their organization and communities alike (Kaufman, 1992; 1998; Popcorn, 1991).

Our Changing World—Expanding  
Kirkpatrick's Model to Include Societal  
Value-added

Developers and evaluators of training and educational programs must be cognizant of the operational realities of the external world. Since the introduction of Kirkpatrick's evaluation framework in 1959, world economies have been changing at rapid speeds. As Toffler (1990) and Drucker (1993) have suggested, the focus of our

economy has been transformed in recent decades from money and things to information. Additionally, organizations in the private and public sectors must address new challenges by professional, technical, and social communities—demonstrating how they add value to society (Kaufman, 1998). Faith Popcorn (1991) has even suggested that “doing societal good is no longer a corporate option, but a must.” And organizations have responded by transforming their focuses in such a manner not only to survive in the new paradigm, but to thrive in it (Pave & Krausz, 1995; Kaufman, Watkins, Triner & Stith, 1998).

While many aspects of performance improvement processes have been transformed in response to these changes in the organization’s internal and external environments, the Kirkpatrick framework—as commonly used in organizational training and educational programs—has not undergone in application the transformation-continuous improvement common in the field. The Kirkpatrick framework is still utilized (or under-utilized) in its initial form—one which does not facilitate a focus on contributions to our communities and external clients. In response to this Kaufman & Keller (1994) introduced an expanded framework that addresses today’s organizational paradigm, one that includes societal value-added as well as a focus on continuous improvement rather than summative evaluation. The Kirkpatrick Plus framework (see

Table 1) relates Kirkpatrick’s initial framework to the Organizational Elements Model (OEM) (Kaufman, 1992; 1998).

The Kirkpatrick Plus framework offers a modified and expanded model that relates Kirkpatrick’s evaluation framework to Kaufman’s OEM for needs assessment and system planning. Differentiating the Kirkpatrick Plus model from the original Kirkpatrick model, as well as Phillips (1996) adaptation of the model, is the expansion of the framework to include explicit examination of organizational and societal results.

The OEM identifies for any organization the necessary linkages and alignments of its Inputs (resources), processes (methods and techniques), products (building block products or micro results), outputs (deliverables to clients or macro results), and outcomes (external client and societal value-added or mega results). While the OEM has provided a model for strategic planning and needs assessment for nearly three decades, the utilization of the model for proactive evaluation (or better, continuous improvement) has only become more widely accepted in recent years. As noted in Table 1, while there are many areas of compatibility, the Kirkpatrick framework and the OEM are not analogous on a one-to-one relationship.

An additional modification to the initial evaluation framework is the separation of process and resource evaluation at Kirkpatrick’s level 1. While remaining aligned with Kirkpatrick’s

Table 1 □ A comparison of Kirkpatrick’s four-level evaluation framework and the five levels of critical organizational concern and evaluation (Kaufman, Keller and Watkins 1995).

<i>Kirkpatrick's levels</i>	<i>Organizational questions/focus</i>	<i>Compatibility</i>	<i>Suggested levels*</i>
(Not addressed)	Societal contributions (Level 5 evaluation)	Missing from Kirkpatrick’s framework	Mega
4. Results	Organizational payoffs	Yes	Macro
3. Behavior (performance)	Individual or small group payoffs	Yes	Micro
2. Learning (acquisition)	Individual or small group payoffs	Yes	Micro
1. Reaction	a. Process acceptability and efficiency	Partial; efficiency and quality are missing from Kirkpatrick’s framework	Process
	b. Resource availability and quality		Input

\* based on Kaufman’s Organizational Elements Model (1992; 1998).

level 1 evaluation, the modified framework enables evaluators to separate the quality and availability of resources from their efficient use and impact during the training process. The Kirkpatrick Plus model thereby encourages improvement professionals to ask two important questions in a level 1 evaluation: (a) Are books, computers, reference materials, and other resources readily available and of the quality required for training and/or educational experience? and (b) Are the training and/or educational processes efficient and acceptable? This change at level 1 of the initial framework is intended to assist improvement professionals and evaluators in their efforts to go beyond the conventional "smiley sheet" and instead provide managers and decision makers with the valuable information necessary for continuously improving organization training and education efforts. The importance of this separation can be derived from the basic concepts underlying quality management and is commonly left out in conventional applications of Kirkpatrick's initial model.

Kirkpatrick has suggested that training professionals' difficulty of evaluating at levels 3 and 4 makes changes to the framework irrelevant (personal correspondence, 1996). This response to infrequent implementation of upper level evaluations (Kirkpatrick levels 3 and 4) suggests several possible explanations, two of which are (a) the framework levels are not responsive to the essential questions of today's managers, or (b) performance improvement professionals do not have the skills, knowledge, or motivation to evaluate at all levels. While it may be convenient to associate the infrequent implementation of level 3 and 4 evaluation today with trainer (in)competence, we suggest that the lack of application is more likely because of missing linkages of the Kirkpatrick framework to necessary information that managers require in making responsible decisions and demonstrating the adding of value to the organization and community.

Among American Society for Training and Development's Benchmarking Forum, results of a question about training evaluation yielded the following information: 92% of courses are evaluated at level 1; 34% evaluated at level 2, 11%

evaluated at level 3; and 2% evaluated at level 4 (American Society for Training and Development [ASTD], no date). As a possible explanation to this trend, Phillips (1996) has suggested that both the value and the complexity of information increase as an evaluator progresses through the levels or steps of the Kirkpatrick framework, thereby decreasing the frequency of use. But does the utility of evaluation end with the *results* level as defined by Kirkpatrick? And does the complexity of evaluation always increase in relation to the value of information? We suggest that the answer to both of these questions is no.

### Going Proactive

The Kirkpatrick framework is based in a paradigm that emphasizes evaluation as a follow-on process to the design, development and implementation of training or educational programs. Though this approach may have produced success in the past for many managers and their organizations, in today's age of rapid prototyping, quality management, reengineering, and continuous improvement, this paradigm is likely incomplete and insufficient. By relating the Kirkpatrick framework to the Organizational Elements Model which has its roots in societal-value-added, the Kirkpatrick Plus framework offers a unique proactive advantage by moving evaluation from the "back end" of a training or educational project to an integral part of all steps throughout the process. The Kirkpatrick Plus framework, utilizing the levels of results identified in the OEM, can be an essential tool for organizations (including training departments and educational institutions) to align their evaluation initiatives with their strategic planning, needs assessment, intervention development, and conventional evaluation processes.

While the differences in proactive and reactive performance improvement are not always easily identifiable, Figures 1 and 2 show how the processes are likely to be distinct when applied with reference to the Organizational Elements Model. The proactive approach to performance improvement ideally begins with the identification of contributions to external clients and soci-

Figure 1 □ A Proactive Approach to Planning and Evaluating (Kaufman, 1998)

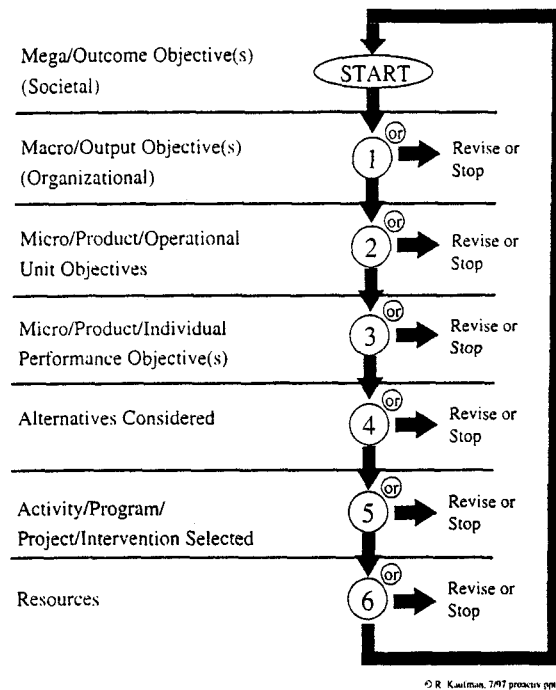
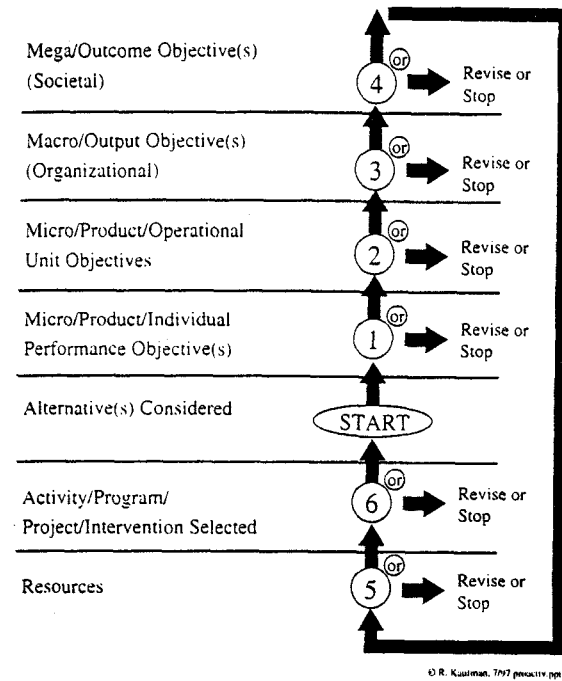


Figure 2 □ A Reactive Approach to Planning and Evaluating (Kaufman, 1998)



ety (or mega-level results) and then “rolls-down” from those results to be achieved to determine the performance intervention(s) and resource requirements. The reactive approach to performance improvement begins where the proactive approach ends, with the performance intervention desired and resources available.

This model of proactive development (see Figure 1) is actually the reverse of the Kirkpatrick evaluation approach. Evaluation at the Kirkpatrick levels 1 and 2 *begins* with the intervention (or instructional material) and tests a series of assumptions . . . beginning with the assumption that the intervention brought about some excitement or level of knowledge acquisition, or both. The second common assumption is that the knowledge will lead to a change in behaviors, which is evaluated at Kirkpatrick level 3. The next assumption is that changed behaviors will produce desired results and this is evaluated at level 4. Unfortunately, very few organizations—approximately 2% (ASTD, no date)—evaluate at level 4. A final assumption of the Kirkpatrick evaluation model is that desired results are synonymous with required results (i.e., that desired results offer some benefit to the

organization and society-community). By using the Kirkpatrick Plus framework to triangulate both *hard* or independently verifiable data and *soft* or perceptual data, evaluators can better ensure that the requirements derived from their activities are grounded in external pay-off, rather than “preferred solutions,” as is common when one begins an evaluation at level 1 of Kirkpatrick’s initial model.

By adopting a proactive framework for evaluation that aligns the evaluation with the planning process, these assumptions can be addressed. By beginning an evaluation or continuous improvement effort at the Kirkpatrick Plus level 5, and identifying criteria at each level of the framework before designing and developing the interventions, alignment and linkages are better assured.

### Criteria

The Kirkpatrick Plus framework, as an evaluation-continuous improvement framework, requires that criteria be developed prior to or during the process. Unfortunately, too often these evaluation criteria are developed after the

training or educational program has been developed and implemented. While this process of development may seem logical and convenient at first (many evaluators commonly even refer to it as a "logic model" approach), it will likely lead to disappointment. More practical is a proactive set of performance measures by which program success is determined prior to the development of the instruction, training, or intervention. The development of evaluation criteria as part of initial design, supported by many instructional systems design models, assures a developmental focus on the required skills, knowledge, attitudes, and accomplishments required to meet the criteria. During a needs assessment, commonly the first step of an instructional design model (e.g., Dick & Carey, 1989; Rothwell & Kazanas, 1992), the performance technologist-instructional designer has an opportunity to assure the alignment or intervention development with organizational and community-society value-added through the utilization of appropriate evaluation criteria. This approach—identifying the useful requirements before any intervention—is more efficient and effective than having an otherwise well-designed intervention fail and then trying to trace down the reason for the failure.

The development of evaluation criteria, though, is often complicated by the utilization of *not quite measurable objectives* (NQMOs) in the instructional design process. Common in most organizations, NQMOs identify only the direction without the rigorous indicators for accomplishment (cf. Gilbert & Gilbert, 1989). An

example of an NQMO can be found in a recent legal debate between the Miccosukee Indian tribe and the state of Florida concerning water quality.

While current legislation calls for "no degradation of flora or fauna," the Miccosukee are calling for a strict numerical standard by which to hold polluters accountable (Muellner, 1998). These NQMOs present a challenge to the developer of the evaluation methods and vehicles because their lack of rigor and measurability on an interval or ratio scale fails to provide a clear linkage between: (a) the performance problem (gap in results) the intervention or educational program was intended to close; (b) the performance to be evaluated; (c) the intended value-added to the organization; and (d) the value-added to external clients and society.

#### Difficulty of Upper Level Evaluations

Philips (1996) offers his own expansion of the Kirkpatrick framework. Phillips includes return-on-investment (ROI) analysis as a fifth level and suggests that this level offers the most valuable information but is the most difficult to assess. We suggest that ROI analysis (or costs-consequences analysis) can actually be utilized to compare the costs of a training or educational program to the evaluated results and consequences—value-added—for any intervention at any level of the Kirkpatrick or Kirkpatrick Plus frameworks (Kaufman, 1998; Kaufman et al., 1995; Kaufman & Keller, 1994; Muir, Watkins, Kaufman, & Leigh, 1998).

Table 2 □ Elements of a Costs-Consequences Analysis (Based in part on Kaufman, 1992; 1998; Kaufman, Watkins, and Sims, 1997; Kaufman, Keller, and Watkins, 1995.)

Type of cost-results analysis				Level of the Organizational Elements Model	Relationship to Kirkpatrick Plus*
Cost-utility	↔	relates costs to	↔	Outcomes	Level 5
Cost-benefit	↔	relates costs to	↔	Outputs	Level 4
Cost-effectiveness	↔	relates costs to	↔	Products	Levels 2 and 3
Cost-efficiency	↔	relates costs to	↔	Processes	Level 1b
Auditing	↔	relates costs to	↔	Inputs	Level 1a

\*From Table 1

Utilizing the Costs-Consequences Analysis model described in Table 2 as a combination of ROI analyses at each level of the Kirkpatrick Plus framework, an organization can perform a systemic, coarse-grained analysis that compares *costs* with *results*. Several levels of analysis can be done in a costs-consequences analysis (e.g., relating the costs of an intervention to the number of students who were satisfied with the efficiency and acceptability of the training, the educational processes, or both; relating the costs to the value-added to the organization and its clients; and relating the costs of an intervention to the value-added for external clients and society). Our experiences with evaluations and costs-consequences analyses at the higher levels of the Kirkpatrick Plus framework (levels 3, 4, and 5) do not coincide with the suggestions of Philips that they are difficult to assess. Kaufman, Watkins & Sims (1997) and Hall, Sprague & Watkins (1995) describe an initial application of a costs-consequences analysis within a state's department of labor. The project tracked both the costs and results of individual participation in the department's Job Training Partnership Act (JTPA) programs. By not requiring the detailed analysis often affiliated with auditing or accounting, the coarse-grain approximations concluded from the study provided managers with only the necessary information for making responsive and responsible decisions.

We are, however, in agreement with Philips in his suggestion that the value of the information attained increases as an organization evaluates at higher levels. In fact, we argue that without the linking to the "plus" (the external and societal consequences) the whole organization and associated interventions are at risk (Kaufman, 1997; Kaufman & Watkins, 1996). The Costs Consequences Analysis model offers organizations a coarse-grain approximation of the ROI that is only at the level of detail required by decisions without the extensive itemized accounts that often fill reports.

Obtaining useful metrics for "opportunity costs" will be essential in the further development of the Kirkpatrick Plus framework. In other words, what could have been the organizational and societal value-added if training, educational resources, or both, had been

expended in other areas of the organization? For example, if an organization did not spend funds on additional diversity training, what else could that same money have delivered in terms of value-added to internal and external clients? These alternative investments play an essential role in determining the cost of an intervention. Of additional interest is the "internal/organizational" costs-consequences of alternative interventions such as modes of delivering training. For example, one could ask, "Instead of investing in classroom-based computer-assisted instruction, what would be the cost consequences of distance learning as a delivery vehicle for learners at remote or home sites?"

A common first step in conducting a costs-consequences analysis is to determine the costs of each program (preferably analysis data are available per student, but often this is not the case). Several basic areas from which costs data can be derived include learner costs, instructors costs, instructional development costs, facilities costs, as well as the opportunity costs of each (Muir et al., 1998). For forecasting, it is best to amortize the costs of the training program over its lifetime. Programs that might be expensive for the short haul can turn out to be comparatively inexpensive over the long haul. A common misconception is that the cost of training development is the most expensive aspect of the training program. Studies have shown, however, that when learner costs are factored in, they often represent more than 80% of the cost of the training program (Foshay, in press).

The variety of data sources available with regard to the consequences (or results) derived from any training or educational intervention can seem overwhelming. The OEM's three levels of results (mega/societal, macro/organizational, and micro/individual or small group) provide a pragmatic framework for structuring the consequences to be utilized in a costs-consequences analysis. By utilizing the three levels of results along with a determination of direct and indirect results of training, the costs-consequences analysis can provide coarse-grain approximations that correspond with Kirkpatrick's levels 2, 3, and 4 as well as the fifth level of Kirkpatrick Plus.

## Summary

While organizations have found that the Kirkpatrick framework for evaluation has met their requirements in the past, the new social and economic realities suggest that evaluators and program-organizational developers reexamine their role and their potential contributions within organizations, and the utility of the models they apply. Given the increasing importance of rigorously demonstrating societal value-added, the Kirkpatrick Plus framework addresses the concerns of communities and organizations, while additionally offering evaluators the opportunity to align their processes with all that an organization uses, does, produces, and delivers, as well as the value it adds to society and external clients. □

Ryan Watkins, Doug Leigh, and Roger Kaufman are with the Office for Needs Assessment and Planning, Florida State University at Tallahassee (<http://onap.fsu.edu/onap>).

Rob Foshay is at TRO Learning, Inc., Hoffman Estates, IL.

## REFERENCES

- American Society for Training and Development (ASTD). (no date). Percent of courses evaluated at Kirkpatrick levels [online]. Available: <http://www.astd.org/who/research/benchmark/96stats/graph15.gif> [March 30, 1998].
- Dick, W., & Carey, L. (1989). *The systematic design of instruction* (3rd ed.). Glenview, IL: Scott Foresman & Co.
- Drucker, P.F. (1993) *Post-capitalist society*. New York: HarperBusiness.
- Foshay, R. (in press). An examination of evaluation research.
- Gilbert, T., & Gilbert, M. (1989: January). Performance engineering: Making human productivity a science. *Performance Improvement*.
- Hall, J., Sprague, D., & Watkins, R. (1995; August). "Florida's job training programs: What is the return on taxpayers' investment?" Tallahassee, FL: Florida TaxWatch Incorporated.
- Kaufman, R. (1992). *Strategic planning plus: An organizational guide*. Newbury Park, CA: Sage (Revised).
- Kaufman, R. (May-June, 1997). Avoiding the "dumbing down" of human performance improvement. *Performance Improvement*.
- Kaufman, R. (1998). *Strategic thinking: A guide to identifying and solving problems—Revised*. Arlington, VA & Washington, DC: Jointly published by the American Society for Training and Development and the International Society for Performance Improvement.
- Kaufman, R., & Keller, J. (1994: Winter). Levels of evaluation: Beyond Kirkpatrick. *Human Resources Quarterly*, 5(4).
- Kaufman, R., Keller, J., & Watkins, R. (1995). What works and what doesn't: Evaluation beyond Kirkpatrick. *Performance and Instruction* 35(2), 8–12.
- Kaufman, R., & Watkins, R. (1996: Spring). Costs-consequences analysis. *HRD Quarterly*.
- Kaufman, R., Watkins, R., & Sims, L. (1997). Cost-consequences analysis: A case study. *Performance Improvement Quarterly* 10(2).
- Kaufman, R., Watkins, R., Triner, D., & Stith, M. (1998: summer). The changing corporate mind: Organizations, vision, missions, purposes, and indicators on the move toward societal payoffs. *Performance Improvement Quarterly*.
- Kirkpatrick, D.L. (1959a). Techniques for evaluating training programs. *Journal of ASTD*, 13(11), 3–9.
- Kirkpatrick, D.L. (1959b). Techniques for evaluating training programs: Part 2—Learning. *Journal of ASTD*, 13(12), 21–26.
- Kirkpatrick, D.L. (1960a). Techniques for evaluating training programs: Part 3—Behavior. *Journal of ASTD*, 14(1), 13–18.
- Kirkpatrick, D.L. (1960b). Techniques for evaluating training programs: Part 4—Results. *Journal of ASTD*, 14(2), 28–32.
- Kirkpatrick, D.L. (1994). *Evaluating training programs: The four levels*. San Francisco: Berrett-Koehler.
- Kirkpatrick, D.L. (1996). Evaluation. In R.L. Craig, & L.R. Bittel (Eds.), *Training & Development Handbook*. American Society for Training and Development, New York: McGraw-Hill Book Co.
- Kirkpatrick, D.L. (1996: Spring). Invited reaction: Reaction to Holton article. *Performance Improvement Quarterly*, 7(1), 23–29.
- Muellner, A. (Reporter). (1998, March 18). *Morning Edition*. Upper Marlboro, Maryland: National Public Radio.
- Muir, M., Watkins, R., Kaufman, R., & Leigh, D. (June, 1998). Costs-consequences analysis: A primer. *Performance Improvement*.
- Pava, M.N., & Krausz, J. (1995). *Corporate responsibility and financial performance: The paradox of social cost*. Westport, CN: Quorum Books.
- Phillips, J. (1996). Measuring the Results of Training. In R.L. Craig, & L.R. Bittel (Eds.), *Training & Development Handbook*. American Society for Training and Development, New York: McGraw-Hill Book Co.
- Popcorn, F. (1991). *The Popcorn report*. New York: Doubleday.
- Rothwell, W.J., & Kazanas, H.C. (1992). *Mastering the instructional design process: A systematic approach*. San Francisco: Jossey-Bass Publishers.
- Toffler, A. (1990). *Powershift: Knowledge, wealth and violence at the edge of the 21st century*. New York: Bantam Books.