Program and Project Evaluation

16

J. Michael Spector

Abstract

Educational technology research and development nearly always involves an intervention of some kind aimed at solving a problem or improving a situation pertaining to learning and instruction. Those involved—the stakeholders—naturally want to know whether the problem was solved and/or the extent to which the situation was improved. Attributing any outcomes to the intervention is not as easy as it may appear, as many factors are typically involved, beyond just the technology involved. This chapter describes a holistic approach to educational technology project and program evaluation. The emphasis is on evaluating the entire process from needs assessment through design, development, deployment, and support with particular attention to evaluating every aspect of the process so as to increase the likelihood of successful technology integration. The use of a logic model to organize evaluation as well as research is described.

Keywords

Confirmatory evaluation • Formative evaluation • Logic model • Needs assessment • Program evaluation • Summative evaluation • Theory of change

Introduction

There are of course many different kinds of research. Typically, research is aimed at determining causes and/or explanations for unusual or unexplained phenomena or at making predictions about what might happen when certain situations occur. Explaining what happened or might happen (and why) as a result of particular educational policies and practices and instructional interventions represents an important kind of research not specifically addressed in previous editions of this *Handbook*. Such applied research is typically called program or project evaluation. Critical questions that inform program and project evaluation include the following:

1. To what extent were the goals and objectives of the program or project met or being met?
2. Did the implementation happen as planned?
3. Was adequate preparation and training provided?
4. Was the design clearly aimed at the problem identified at the outset?

One can of course imagine other questions, some of which will arise in the course of this chapter. First, however, it is necessary to distinguish products, policies, practices, programs, and projects. The latter two are easily distinguished. A project is typically aimed at addressing a particular problem situation by introducing something new or different, which can be called an intervention. As a consequence, a project has a goal and objectives, a beginning (could be the start of the needs assessment but is more typically the start of the development of the intervention), and an ending (typically a short time after the intervention has been deployed

J.M. Spector (✉)
Department of Learning Technologies, College of Information, The University of North Texas, 3940 N. Elm, Suite G150, Denton, TX 76207, USA
e-mail: mike.spector@unt.edu

J.M. Spector et al. (eds.), *Handbook of Research on Educational Communications and Technology*,
DOI 10.1007/978-1-4614-3185-5_16,

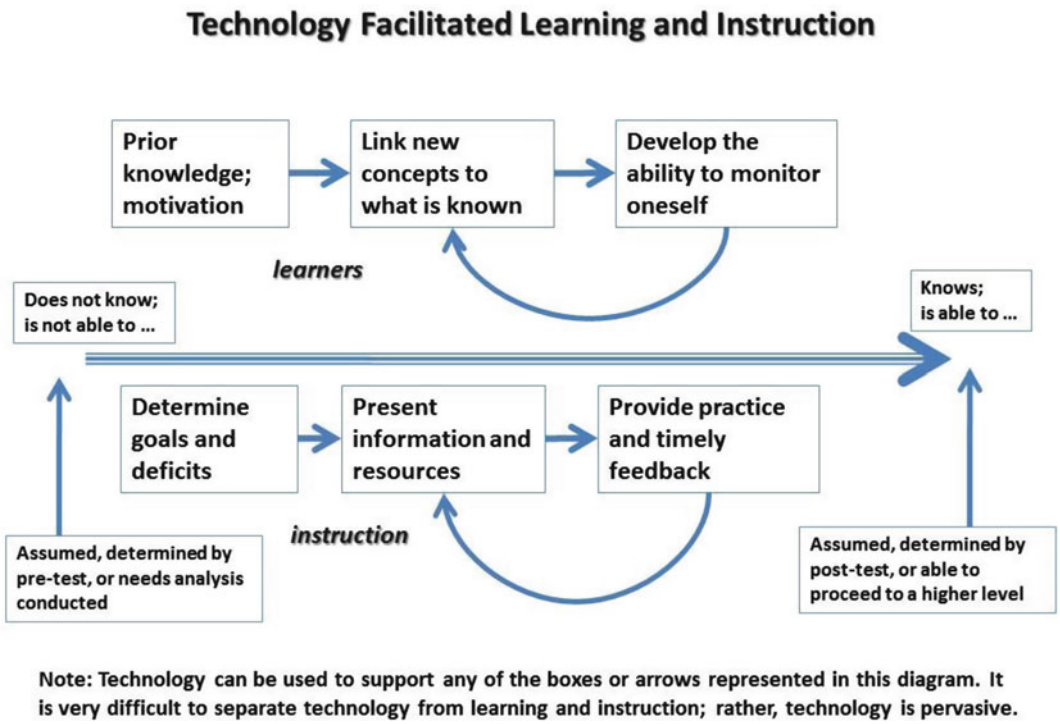

Fig. 16.1 Research on technology in learning and instruction (adapted from Spector, 2010)

and is being used regularly). A program shares many of the same attributes, but unlike a project a program is typically intended to continue in use and evolve over time, so the period of time associated with a program evaluation is longer than that associated with a project evaluation, and changes in the situation surrounding the program need to be actively considered as the program is periodically reviewed. Confirmatory evaluation is a kind of evaluation appropriate for determining if the original assumptions and problem situation are still pertinent to an ongoing program or a long-term project. Projects often mark the initiation of a new program, so there is a close association between projects and programs. Moreover, programs and projects typically alter practice (how those involved conduct their affairs and accomplish specific tasks), and they may involve the introduction of products, policies, and procedures to guide practice and the use of new technologies. This chapter is not focused on the evaluation of policies, procedures, or products, although some of the same principles and techniques are likely to be pertinent in those contexts as well.

Next, there are the notions of learning, instruction, and technology to consider. A technology involves the disciplined or systematic application of knowledge to solve a particular problem or achieve a specific goal (see Spector, 2012). Examples of educational technologies include online discussion forums, animated models, interactive simulations, checklists for procedures, mnemonic memory aids, and much more. Learning involves recognizable and persisting changes in an individual's (or organization's) abilities, attitudes, beliefs, knowledge, or skills (Spector, 2012). Instruction is that which supports or facilitates learning (Gagné, 1985; Spector, 2012). Based on these common definitions, it is obvious that there are close connections between learning, instruction, and technology, as depicted in Fig. 16.1. Just as technology can support any of the boxes in Fig. 16.1, each of those boxes can and should have an associated evaluation activity. Evaluation, properly understood, pervades educational practice and is essential for understanding the impact of projects and programs.

One can conceivably ask many different kinds of questions about a variety of technologies used to support and facilitate learning and instruction. Many of these questions involve the characteristics of scientifically based research, including the notions of baseline studies, comparison and control groups, analyses of significance and effect size, growth curve modelling, and so on (Cronback, 1989; Nagel, 1994; Scriven, 1960, 1994; Spector, 2010; Suchman, 1967; Suppes, 1978).

The history of educational research is rich and diverse, with more than 100 years of empirical investigations (Aldrich, 2002; Knox, 1971; Langemann, 2000; Suppes, 1978). In spite of such a large body of evidence, there is little evidence that the many educational technology innovations that have been introduced in the last 100 years have had a significant impact on learning (Langemann, 2000; Russell, 2001; Suppes, 1978). Several explanations for this unusual

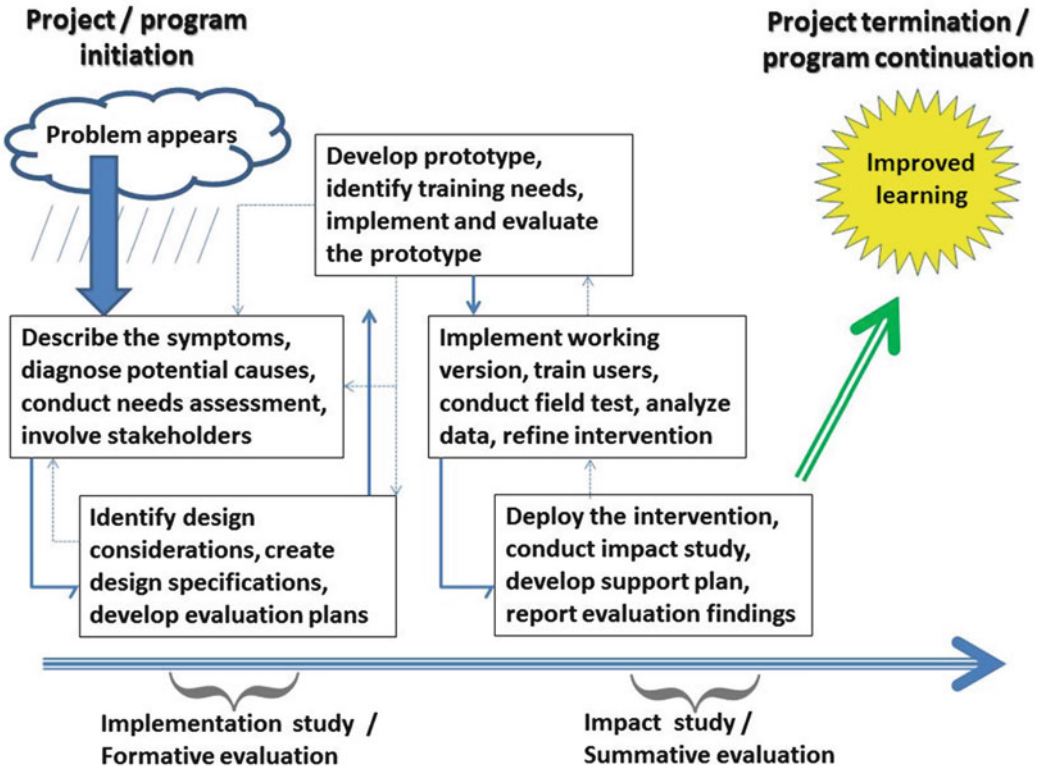

Fig. 16.2 A representation of projects and programs

finding have been provided, many of which focus on the inadequate kind of evaluation research conducted in association with the integration of new technologies into learning and instruction (Langemann, 2000; Russell, 2001; Suppes, 1978). A major inadequacy of prior research on the impact of educational projects and programs involves a tendency to examine superficial indicators of impact without examining the quality of the implementation and training associated with an intervention. Typically, project evaluations have focused on three indicators of success: (a) Did the implementation stay within budget? (b) Did the implementation occur on schedule? and (c) Did the implementation meet the design specifications? Such evaluations are summative in nature—that is to say that they do not provide any information or input that will improve the development of the effort while it is under way. Moreover, associated with such a summative evaluation there might be a research effort that looks at learning outcomes before and after the intervention was developed and deployed. The analysis of the before and after research data often indicates that there is little impact on learning (Russell, 2001; Suppes, 1978), although the three project indicators may reflect success. What is one to think?

The conclusion that will be elaborated in what follows is that a summative evaluation is not adequate and serves little real purpose without the support of a thorough formative evaluation along with a confirmatory evaluation for longer term efforts. Project and program evaluation can and should be aimed at the entire life cycle of the effort and be designed to ensure that the development effort will not only meet the design specifications but also address and solve the indicated problematic situation (see Fig. 16.2). Formative evaluation that is intended to improve an intervention as it is being designed and developed is required; this notion goes to the heart of what project and program evaluation is really about—one does evaluation to help ensure that time and money are not wasted, which means that evaluation must begin early and continue as the effort progresses. One can make the same argument with regard to assessing student learning—the proper emphasis is on helping to improve student learning and not merely on reporting what learning seems to have occurred at the end of a sequence of learning activities.

Confirmatory evaluation is worth emphasizing at this point as it helps distinguish projects from programs and reinforces the formative and summative nature of evaluation. Confirmatory evaluation is aimed at reexamining the problematic situation after an extended period of time, often after a project has been completed and a program has been under way for some time; long-term projects often revisit the needs assessment and requirements analysis phase of a project to make sure that essential aspects of the problem are still the same as originally identified. Confirmatory evaluation is conducted to ensure that the right problem is being addressed and solved. More specifically, confirmatory evaluation involves a systematic program analysis that is aimed at

attributing effects to causes as well as reexamining assumptions and the original problem situation; this is important in explaining significant effects as well as the lack of significant effects (Reynolds, 1998).

Logic Models and Program/Project Evaluation

By way of summary, when conducting applied or development research, one may have a new educational technology or system that one believes will be beneficial in some way. This situation is a prime target for research and inquiry. One kind of inquiry often associated with development research is program (or project) evaluation. The basic questions are whether and to what extent an intervention (e.g., an innovative technology or new learning environment or educational system) achieves its intended aims, and why it succeeded or fell short in some way. The emphasis is not on the technology as a product but rather on its use in promoting learning and/or improving performance. One can imagine two kinds of studies emerging from a program evaluation: a fidelity of implementation study (a kind of formative evaluation) and an impact study (a kind of summative evaluation) (see Fig. 16.2). The notion of a logic model can be used to explain the differences in these two kinds of research and evaluation studies (see Fig. 16.3). A logic model portrays a current situation and the associated problem, the implementation of an intervention intended to address the problem situation, and the projected or the predicted outcomes and benefits of that intervention if and as successfully implemented. A theory of change that would explain why and how the intervention would lead from the problem state to the desired outcomes is normally associated with and depicted in a logic model. The fidelity of implementation study could be structured such that the results of the study reflected degrees of successful implementation (as in high, medium, low, or superior, adequate, or marginal involving such variables as professional development and technology support). Having such data is useful in explaining why and to what extent significant differences were or were not found in outcome variables. For additional detail on such studies, see the chapter by Jennifer Hamilton in this *Handbook*.

The problem description is important as that is the outcome of some kind of analysis typically called a needs assessment.

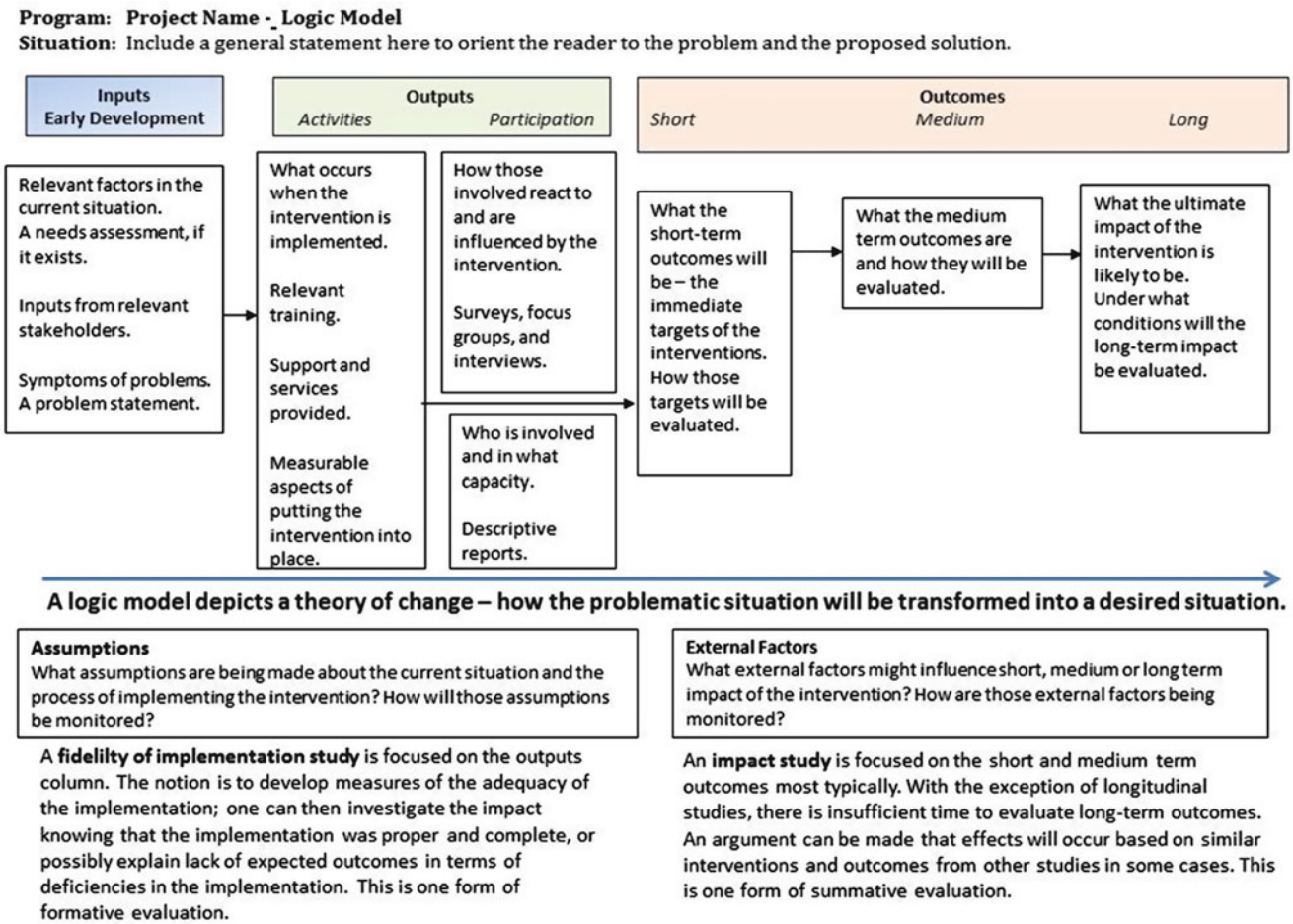

Fig. 16.3 Logic models, fidelity of implementation, and impact studies

The problem indicators become the targets of desired outcomes of the effort, which means that the impact study (summative evaluation) will measure the extent to which the problem situation has improved. However, the responsibility of program evaluation far exceeds simply reporting outcomes, even when resources (time, funds, etc.) are included in the outcomes analysis. The work of program and project evaluators begins with the analysis of the problem situation. Who was involved in the needs assessment? Were all stakeholders involved? Was there a divergence of views? How reliable were the methods and instruments used to collect and analyze needs assessment data? If the problem identification process is flawed, then it is unlikely that the subsequent development effort will produce desired outcomes. Evaluators should observe the early analysis and planning phases and provide formative feedback to help ensure that a comprehensive and high-quality needs assessment drives the effort.

Since nearly all technology implementation efforts involve training users, it is important for evaluators to observe and report outcomes of training plans as well as actual training. Inadequacies in training often result in suboptimal outcomes even when the technology implemented is of high quality. Moreover, planning for change and properly preparing users for a new technology are critical for the successful diffusion of an innovation within an organization (Ellsworth, 2000; Rogers, 2003). Again program evaluators have a responsibility to report any shortcomings or potential problem areas concerning preparing users for an innovation; this responsibility falls into the category of formative evaluation. A formal analysis of how well the project or program is preparing end users for the effective integration of a new technology is known as a fidelity of implementation study. Such a study is primarily aimed at the inputs and outputs in the logic model (see Fig. 16.3) and the degree to which the intervention can realistically be expected to support the underlying theory of change if it is properly or fully implemented.

A primary obligation of evaluators is to alert the implementation team and management of anything that might jeopardize the desired outcomes of the effort at any time during the entire process of planning and implementing an intervention. However, it is quite rare to find a project or a program that involves evaluators throughout the process in this way. Perhaps this lack of ongoing formative evaluation is another reason that few significant differences are reported even for well-supported educational technology efforts.

Evaluation vs. Research

Why include a chapter on program evaluation in a research handbook? Hopefully the answer is obvious enough at this point. In one sense, an evaluation effort represents one kind of applied research in the sense that an explanation for what has happened (or failed to happen) is developed, especially through the quantitative and qualitative data used in the fidelity of implementation study. Evaluators often use the same tools and methods used by other educational researchers. One difference is the focus of evaluation vs. that of other forms of educational research. Evaluation is focused on decisions made during the planning and implementation of an intervention with the aim of helping to improve the effort so as to produce desired outcomes. Other forms of educational research are focused on answering questions that contribute to a body of knowledge or the development of theories to explain a range of phenomenon (see Popper, 1963). Of course it can and does happen that program and project evaluations inform research about the many phenomena associated with learning and instruction, so the distinction between educational project/program evaluation and education research is not sharp and distinct.

An Example

To suggest how the application of this approach to program evaluation might work, a brief discussion of an invented case is provided here. Obviously other approaches, methods, and instruments are possible to use. This case is meant to emphasize the formative nature of program evaluation as this is the most challenging aspect to implement.

A school district has determined that a significant percentage of its students are at risk. In making this determination, the district used an expanded set of indicators that included some things beyond the control of district personnel (e.g., socioeconomic status, physical and learning disabilities, etc.), and some things that district personnel (both administrators and teachers) believed that they could influence (e.g., absenteeism, test scores, pass/fail rates, graduation rates, discipline cases, etc.). The expanded set of indicators showed many more students at risk than teachers, administrators, and parents had previously imagined. District leadership then initiated focus group discussion with administrators, teachers, parents, and students to determine what problems these groups perceived as most relevant to the situation. Two quite different kinds of problems were mentioned most often: (a) instruction that was not well suited to individual student needs, and (b) lack of easy, real-time access to data that would enable teachers and administrators to be more responsive to individual student needs.

This needs assessment took place over the course of an entire year; it led the district to settle on a solution approach that involved an integrated data system that could provide data on individual students and that could also support differentiated instruction and personalized learning. Since the external evaluator was not involved in the needs assessment, there was an initial concern with regard to confirming the

problem situation as represented by district personnel. As it happened, the district had documented the process quite thoroughly, so the first thing the evaluator did was to review that documentation and confirm that the needs were real and worth addressing through additional focus group discussions with the relevant stakeholders.

Having settled on an approach to resolve the problematic situation, a grant was submitted and funded, a company was hired, and a project team put in place that included an external evaluator. The goals of the project were focused on reducing absenteeism and discipline rates, increasing graduation rates, and improving test scores. These became the outcome indicators in the project's logic model developed by the external evaluator. The theory of change was based on the notion that personalizing learning and making instruction relevant to individual student performance and interest would result in increased student interest and performance, thereby reducing absenteeism and discipline problems while increasing test scores and graduation rates. That theory of change had some support in the published research literature, and it proved convincing to the funding agency. In this case, the impact study (summative evaluation) was easy to construct and implement since the measures were obvious and the data readily available. As it happened, the outcome measures did not provide sufficient basis to say that the intervention made a significant difference. In part, district personnel and the evaluator believed that this was because there was an influx of new students to the district whose first language was not English and because the economic downturn caused some high school students to take part-time jobs leaving them with less time for studies. These facts partially explained the lack of significant impact in terms of the outcome indicators.

During the project, the evaluator collected information from administrators, faculty, students, and parents with regard to training for and use of the new data system. The focus was on the ability of the system to support teachers' needs for timely information and students' needs to have customized learning activities. Prior to the implementation of the system, the evaluator noticed that the company originally hired to provide the system was not responding to teachers' needs for real-time information nor would the system be able to support personalized learning and differentiated instruction without extensive teacher intervention, which was not possible given existing workloads and enrollments. The evaluator recommended requiring the company to comply with the district's requirements or find a company that would. The district followed the evaluator's recommendations and found a company that was able to deliver a system that was responsive to the needs of both students and teachers. Due to a delay, however, the initial training was not as thorough as it could have been. The evaluator documented the time spent on training and problems that teachers and students had with the new system; the evaluator then recommended additional training and support materials, which were developed, but not soon enough to impact the outcome indicators by the time the funded project ended.

Two things are worth noting. First, the evaluator (who is fictitious of course) was able to add to the explanation of lack of significant difference on outcome indicators due to the implementation study that focused on the development process and the training of teachers. In addition, the district did manage to deploy a new system that is having an impact, although that impact was delayed due to the change in software providers. While the project itself could not report significant differences due to personalized learning and an integrated data management system, the project has evolved into a program that is now reporting significant differences, in part thanks to the formative evaluator's recommendation with regard to the software development provider.

Conclusions

This chapter is intended as an introduction to an important area of educational research called program (or project) evaluation. The treatment of program evaluation here was not intended to be comprehensive or describe specific evaluation methods or tools. Rather, the intent was to stress the significance of fidelity of implementation studies as they serve to explain the findings in an impact study, and to emphasize the responsibility of evaluators to report potential problem areas during a development effort to the implementation and management team. Readers can find a wealth of information on specific program evaluation methods and tools elsewhere (Louw, 1999; Potter, 2006; Rao & Woolcock, 2003; Rossi, Lipsey, & Freeman, 2004).

References

Aldrich, R. (Ed.). (2002). *A century of education*. London: Routledge/Falmer.

Cronback, L. J. (1989). *Designing evaluations for educational and social programs*. San Francisco, CA: Jossey-Bass.

Ellsworth, J. B. (2000). *Surviving change: A survey of educational change models*. Syracuse, NY: ERIC Clearinghouse.

Gagné, R. M. (1985). *The conditions of learning*. New York, NY: Holt, Rinehart &Winston.

Knox, H. (1971). *A history of educational research in the United States*. Washington, DC: National Institute of Education. Retrieved January 3, 2012, from http://www.eric.ed.gov/PDFS/ED088800.pdf.

Langemann, E. C. (2000). *An elusive science: The troubling history of education research*. Chicago, IL: The University of Chicago Press.

Louw, J. (1999). Improving practice through evaluation. In D. Donald, A. Dawes, & J. Louw (Eds.), *Addressing childhood adversity* (pp. 66–73). Cape Town: David Phillip.

*Nagel, E. (1994). Introduction: Philosophy in educational research. In S. R. Sharma (Ed.), *Encyclopedia of modern educational research* (pp. 1–16). New Delhi: Anmol Publications.

Popper, K. (1963). *Conjectures and refutations: The growth of scientific knowledge*. London: Routledge.

Potter, C. (2006). Program evaluation. In M. Terre Blance, K. Durrheim, & D. Painter (Eds.), *Research in practice: Applied methods for the social sciences* (2nd ed., pp. 410–428). Cape Town: UCT Press.

Rao, V., & Woolcock, M. (2003). Integrating qualitative and quantitative approaches in program evaluation. In F. Bourguignon & L. Pereira da Silva (Eds.), *The impact of economic policies on poverty and income distribution: Evaluation techniques and tools* (pp. 165–190). Oxford: Oxford University Press.

Reynolds, A. J. (1998). Confirmatory program evaluation: A method for strengthening causal inference. *American Journal of Evaluation, 19*(2), 203–221.

Rogers, E. M. (2003). *Diffusion of innovations* (5th ed.). New York, NY: Free Press.

Rossi, P., Lipsey, M. W., & Freeman, H. E. (2004). *Evaluation: A systematic approach* (7th ed.). Thousand Oaks, CA: Sage.

Russell, T. L. (2001). *The no significant difference phenomenon: A comparative research annotated bibliography on technology for distance education*. Montgomery, AL: The International Distance Education Certification Center.

*Scriven, M. (1960). The methodology of educational research. *Review of Educational Research, 30*(5), 422–429.

Scriven, M. (1994). The fine line between evaluation and explanation. *Evaluation Practice, 15*, 75–77.

Spector, J. M. (2010). Mental representations and their analysis: An epistemological perspective. In D. Ifenthaler, P. Pirnay-dummer, & N. M. Seel (Eds.), *Computer-based diagnostics and systematic analysis of knowledge* (pp. 17–40). New York, NY: Springer.

*Spector, J. M. (2012). *Foundations of educational technology: Integrative approaches and interdisciplinary perspectives*. New York, NY: Routledge.

*Suchman, E. A. (1967). *Evaluation research: Principles and practice in public service and social action programs*. New York, NY: Russell Sage Foundation.

*Suppes, P. (1978). *Impact of research on education: Some case studies*. Washington, DC: National Academy of Education.