

Evaluation of Adult Education and Training Programs

G McNamara, P Joyce, and J O'Hara, Dublin City University, Dublin, Republic of Ireland

© 2010 Elsevier Ltd. All rights reserved.

Glossary

Competences – The integration of knowledge, skills, and attitudes concerned primarily with developing the ability to acquire further knowledge, rather than simply the possession of a particular amount of current knowledge.

Evaluation – The systemic assessment of an object's worth, probity, feasibility, safety, significance, and/or equity.

Formative evaluation – Evaluation conducted during the development of a program.

Program evaluation – The use of social research procedures to systematically investigate the effectiveness of social intervention programs such as education and training.

Quality – The ongoing and continuous analysis of the provision and outcomes of programs.

Summative evaluation – Retrospective assessments of completed or established programs.

Introduction

Evaluation occupies an increasingly major place in public and private decision-making processes, serving many political functions such as analysis of spending, allocation of funds, and provision of accountability. On the other hand evaluation is also increasingly concerned with learning, with empowering program participants and enabling social transformation.

Program evaluation is the use of social research procedures to systematically investigate the effectiveness of social intervention programs such as education and training (Stufflebeam and Shinkfield, 2007). At one level the demand for an appropriately skilled workforce in an evolving global economy has made the evaluation of adult education and training programs a high priority but one which seemed relatively straightforward. However ideological, curricular, and political rethinking has fundamentally undermined this seeming simplicity. The conceptual framework underpinning theories of adult education and training has become deeply contested. In terms of curricular structure the traditional focus on behavioral outcomes has been replaced by the broader concept of competences.

Equally, increased demands for value for money in public services have resulted in concerns about accountability and quality moving to the fore. Evaluation has also progressed dramatically and is now no longer simply concerned primarily with measurable outcomes but with process, stakeholder roles, values, and quality.

In this article, the contextual issues to be considered in the evaluation of adult education and training programs are outlined. Key questions which need to be answered whichever approach is undertaken are presented. Evaluation models and approaches are summarized and two particular models are highlighted as popular but contrasting approaches to evaluating adult education and training programs. Finally a greater emphasis on self-evaluation is suggested as an approach worth considering for the twenty-first century in the field of adult education and training.

The Changing Face of Evaluation in Adult Education and Training

Evaluation theory and practice in adult education and training has, in recent times, undergone significant development. This development reflects not only evolving concepts of evaluation but equally dramatic changes in the philosophy and curriculum of adult education and training. For example, education and training in nursing and medicine have undergone significant reform in the past 10–15 years. Continuing with a strong practice-based focus it now also has an explicit focus on problem-based learning, change management, and policy development. The focus has also changed from one of instruction, for example, an apprenticeship model, to a learning paradigm, where the emphasis is on student-centered learning. The role of the adult learner is one of active participant who sees a relevance to learning which can be applied to practice.

The purposes and ideology of much adult education and training have become deeply contested. Traditionally, adult education and training had been seen as primarily concerned with issues such as skills, labor market productivity, and economic development. More recently, post-modernist thinking has moved the focus to issues of social transformation concerned primarily with personal and societal development as key purposes in adult education. Particularly influential have been constructivist theories

reminding us that knowledge, while individually constructed, requires collective and collaborative interaction as a key element of learning. Thus, while for certain purposes objectives-based training in the form of skills development may well be defensible and appropriate, the concept of adult education and training has expanded significantly in line with postmodernist views of epistemology and ontology.

In the light of this changing conceptual focus, the nature and structure of the curriculum in adult education and training has also been reformed. In response to changes in society, and in particular the emergence of the knowledge economy, adult education and training is increasingly concerned with the development of broad competences. Competences integrate knowledge, skills, and attitudes and are concerned primarily with developing the ability to acquire further knowledge rather than being simply about the possession of a particular amount of current knowledge. In consequence, the structure of adult education and training programs has undergone radical change, with an ever greater emphasis on issues of process and less concern with the acquisition of a particular skill set.

Somewhat paradoxically, in tandem with this change in the philosophy and curriculum structure of adult education and training programs, there has been also a parallel increase in policies designed to ensure quality and value for money in the provision of such programs. This is in line with the international neoliberal movement often referred to as new public management which has resulted in greater oversight of programs and initiatives in the public sector and determination to introduce the alleged efficiencies of the free market. Quality is understood to be concerned with the ongoing and continuous analysis of the provision and outcomes of programs (Kells, 1992). It is a broad concept which includes both external requirements for accountability and accreditation such as national standards and benchmarks and internal control of quality within educational institutions and programs.

Evaluation has become dramatically altered as a result of these developments. From its traditional role as a once-off measurement of program outcomes, evaluation is now perceived as an integral part of a continuous cycle of quality assurance which includes program philosophy, curriculum development, definition of quality standards, assessment, strategic planning, and internal and external evaluation. Moreover, since in most cases evaluation, as now constituted, encompasses both accountability and improvement focuses it must be multilevel, capable of responding to different needs and expectations from a variety of audiences.

The above trends have paralleled and are closely linked with changing concepts and practices in the evaluation of adult education and training. Consistent with

developments in other fields there is an increased focus in evaluation on methods which include community or stakeholder input from the beginning of program definition and design. Such methods are designed to help program participants to evaluate themselves and their programs, still with the goal of improving outcomes but also of fostering autonomy and decentralized decision making.

Before the days of competences and quality assurance, the evaluation of adult education and training was primarily concerned with the measurement of traditional behavioral objectives. More recently, evaluation theory and practice has become increasingly defined by a more sophisticated analysis of programs involving the inclusion of stakeholder perceptions and multiple forms of evidence, data, and indicators. There has been a tendency to break away from the classical, objectivist, outcome-based, and performance-orientated concept of evaluation toward a multiplicity of new models. Among these models are responsive evaluation as illumination, ethnographic evaluation, naturalistic evaluation, utilization-focused evaluation, the integrated information development and evaluation model, fourth-generation evaluation, empowerment evaluation, participative evaluation, self-evaluation, and others. In the more recent past, there has been a move from debates between positivists and post-positivists to a dialog between the paradigms leading to a greater emphasis on multimethod approaches.

Therefore, the design of evaluation of adult education and training programs is now more complex than in the past. It has to take into account the changing priorities of the curriculum in such programs, emphasizing key competences at many levels. Evaluation must also be reconceptualized to fit within structures designed not as once-off appraisals but rather continuous cyclical quests for improvement. In design it must reflect these imperatives, and in consequences, educational evaluation theory and practice has moved from simplistic notions of measuring outcomes to more complex concerns with stakeholders' roles and the process of learning.

Designing Evaluation for Adult Education and Training

From the aforementioned discussion it should be clear that the design and conduct of evaluation in the field of adult education and training presents difficulties peculiar to that field. For example, a curriculum based on the complex notions of competences requires standards, indicators, criteria, and appropriate assessment procedures if it is to be coherent. All these features must be evaluated. Since by their nature many competences are in the expressive domain and resistant to traditional notions of measurement, a variety of largely qualitative methodologies will be

applied alongside more traditional modes of testing. If, as likely in the bulk of cases, the evaluation is concerned with both accountability and improvement, and internal and external audiences, the evaluation design must be able to meet all these needs. This may be achievable in theory but recent work in the field by [McNamara and O'Hara \(2004\)](#) suggests that evaluation for accountability and for improvement are not complementary but competing, and that formative goals focused on learning can easily be derailed by the demands for summative judgment. Finally, in the modern discourse of evaluation, the emphasis on negotiation and collaboration and on iterative research processes and methods are often at odds with contractual requirements and realistic resources and timescales.

The increasing complexity of educational evaluation design is illustrated by the definition by [Stufflebeam and Shinkfield \(2007\)](#), which describes evaluation as the systematic assessment of an object's merit, worth, probity, feasibility, safety, significance, and/or equity. This more values-oriented definition is an extension of that put forward by the [Joint Committee on Standards for Educational Evaluation \(1994\)](#) in the USA which defined evaluation as the systematic assessment of the worth or merit of an object. The merit of a program can be judged by examining if it does well what it is supposed to do. Worth refers to a program's combination of excellence and service in an area of clear need within a specified context. In evaluating probity, assessments are made of the program's honesty, integrity, and ethical behavior ([Stufflebeam and Shinkfield, 2007](#)). A good evaluation should provide direction for efficient use of time and resources and be politically viable. Many programs require evaluations that examine safety of facilities, influence, importance, and visibility. Evaluators need to consider the possibility that a program can have far-reaching implications outside the local arena and evaluate its significance over time and in other settings. The criterion of equity argues for equal opportunities for all people and emphasizes freedom for all. It is also concerned with the roles and rights of all legitimate stakeholders including politics, ideology, and imbalances in power relationships. Taking this array of concerns into account in evaluation design is clearly a formidable task.

In the following section we look at two contrasting models of adult education and training evaluation. Kirkpatrick's four-level model is very much in the Tylerian behavioral objectives measurement tradition and is still widely utilized in the evaluation of skills-training programs ([Thackwray, 1998](#)). In contrast [Jacobs' \(2000\)](#) approach is representative of models concerned with stakeholder empowerment, constructivist concepts of knowledge, and social transformation. We begin by explaining the categorization of evaluation into formative and summative and propose key questions which need to be answered whichever approach is undertaken. Finally, the increasing emphasis on self-evaluation is mentioned and it is suggested

that for both conceptual and practical reasons it will become increasingly influential in the field of adult education and training.

Evaluation Models

Designing an evaluation of adult education and training programs involves a number of key steps. Such a structure is important so as not to omit any vital parts of the process. In addition the timing of evaluation is very important and this should be decided in advance of setting up any adult education and training programs.

Formative and Summative Evaluations

The timing of evaluations has been generally differentiated into summative evaluations (retrospective assessments of completed or established programs) and formative evaluations (conducted during the development of a program). Formative evaluations are used to modify and improve a program and this is frequently used to provide feedback to staff while the program is in operation. These evaluations assess and assist with the formation of goals and priorities, provide direction for planning, and guide program management. Information from formative evaluations is directed to improving operations and serves quality assurance purposes. In contrast, summative evaluations are used to prove something, satisfy accountability, or make a judgment about the overall quality of the program. They draw together previously collected information, for example, from formative evaluations. Both formative and summative evaluations are needed in the development of a program ([Rossi et al., 1999](#)). In general, formative evaluation will be dominant in the early stages of a program and summative will take over as the program concludes.

Key Questions

Regardless of model or approach chosen, a number of key questions need to be answered when organizing a program evaluation ([Thackwray, 1998](#)):

- How is evaluation defined?
- What are the functions of the evaluation?
- What are the objects of the evaluation?
- What kinds of information should be collected regarding each object?
- What criteria should be used to judge the merit of an evaluation object?
- Whose interests should be served by the evaluation?
- What methods of enquiry should be used?
- Who should do the evaluation?
- By what standards should the evaluation be judged?
- How and when should the results be presented?

The process of considering these issues and integrating them into evaluative research design is very well exemplified in two evaluation models which have become hugely influential, namely case study evaluation as developed by Robert Stake and utilization-focused evaluation created by Michael Quinn Patton. An interesting range of examples of evaluation practice in complex learning organizations which respond to the above questions is provided in a recent publication of the American Evaluation Association (Braverman *et al.*, 2008).

Kirkpatrick’s Model

Introduced over four decades ago, this model is still used by many for evaluating training and development programs and can be categorized as a results or goal-based evaluation (Figure 1). Initially a four-level model, it was later adapted to include a fifth level to measure return on investment. Each level measures different but complementary aspects of training and development. In essence, Kirkpatrick sought to stimulate those with responsibility for the management of training and development to increase their efforts in evaluating training and development actions. Critics of Kirkpatrick assert that his evaluation process may not always produce genuinely meaningful, long-term results. This model implies that evaluation is a standardized, prepackaged process, which is clearly not always the case.

Reaction or level one is easy to measure and quantify and is interpreted as determining how people feel about the program. Criticisms of this level state that it does not measure any learning which takes place so that moving to level two is important. Recommendations for level two include the use of a before-and-after approach so that learning can be related to the program. Evaluation of behavior or level three of this model is more difficult. It may be possible to appraise performance before and after the program or to have a posttraining appraisal 3 months or more after the program so that participants have an opportunity to put into practice what they learned. The fourth or results level is the most difficult area to evaluate effectively. This level defines results to include an institution’s capacity to learn, change, and develop in line

with its stated agreed objectives. MacDonald *et al.* (2000) used Kirkpatrick’s model to examine the impact of applying adult education principles to training in advanced technology companies. The evaluation explored levels one to four and data was collected by observation notes, post module surveys, post assignment surveys, and the participants’ summative evaluation of the program. The authors suggest that the assignment data and summative evaluations confirm a level four on Kirkpatrick’s model was reached on this evaluation.

Jacobs’ Model

However, for many evaluators today, the process as well as the outcomes of modern evaluation has to be meaningful and compelling for the stakeholders if they are to have any useful impact on daily activities in organizations (Figure 2). The epistemic changes in both adult education and training and evaluation theory and practice outlined earlier in this article demand an operational model that permits built-in evaluations with continuous streams of information for both descriptive and evaluative uses. These changes also demand a methodology of evaluation that covers both the technical and social aspects and is comprehensive enough to include the historical, the political, the analytical, and the naturalistic aspects. Such an evaluation will be supported by statistics and stories, quantities and qualities, careful description, and impressions and anecdotes. The current literature on evaluation has many examples of such models. The following based on the work of Jacobs is representative.

Jacobs outlines a ten-stage evaluation model. The opening three stages are concerned with collaborative and negotiated understandings around the context and policy framework of the intervention, the goals of the evaluation, and the identification of and consultation with the principal stakeholders and relevant constituencies. Stages four through seven again involve iterative negotiated agreement on evaluation purposes and criteria, the interests to be served by the evaluation, the aspects of the program to be evaluated, the sources of information to be used, and the evaluation methods most appropriate. Stage eight involves the collection of data from the agreed sources and stage nine involves the analysis and interpretation of the data in the context of the agreed, negotiated criteria. Stage ten involves the dissemination of findings to the stakeholders and constituencies identified at the beginning of the process.

Jacobs stresses that evaluators should employ this negotiated and iterative process, regardless of where the call and resources for the evaluation have come from. She argues that it is central to the appropriate role of the evaluator to recognize the interests and rights of all legitimate stakeholders and to take into account power differentials between groups. Jacobs also advises that

Level 1 Reaction:	What the participants felt about the project or program – the happy sheet?
Level 2 Learning:	Internal validation – where the objectives met?
Level 3 Behavior:	External validation – has training transfer taken place?
Level 4 Results:	Has the project/program made a difference? That is, what has been the impact on the institution?

Figure 1 Kirkpatrick’s four-level model of evaluation.

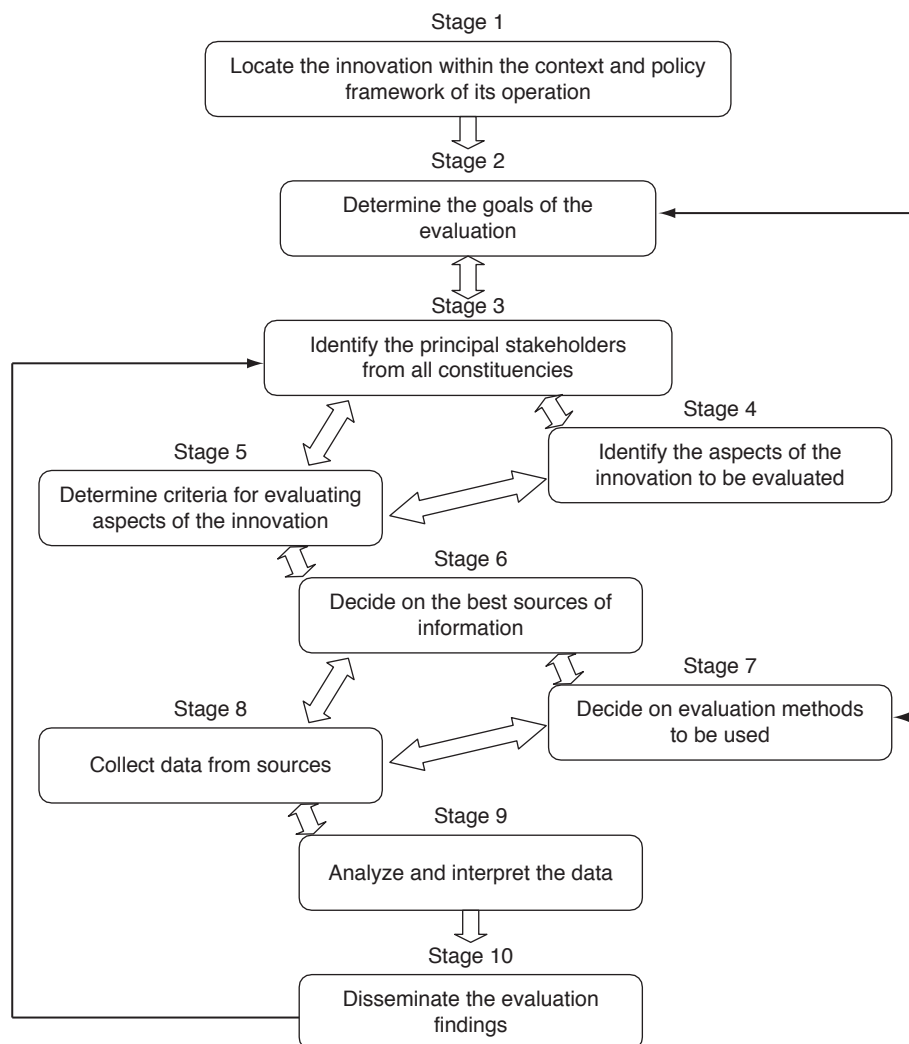


Figure 2 Stages in the process of evaluating educational innovation. Adapted from Jacobs, C. (2000). The evaluation of educational innovation. *Evaluation* 6(3), 261–280, with permission from Sage.

evaluators should give due weight to formative, summative, and illuminative goals and draw from the widest possible range of evidence.

Process-centered models of evaluation such as Jacobs' are widely used in areas of adult and vocational education where the emphasis is on personal and professional development rather than skills acquisition. A good example of collaborative and iterative evaluation in this vein can be found in the work of Bhola (1998).

Self-Evaluation

Despite the developments in evaluation theory and practice outlined so far, many of these models, including more progressive recent ones such as Jacobs', represent a form of evaluation which involves judgments made through the eyes of the external evaluator and the connotation persists of evaluation as an external monitoring of professional practice. The professionalization of evaluation,

the dominance of the contract and terms of reference, and the increasing use of consultants with little knowledge of the field in which they are trying to apply generic research methods are all likely to contribute to the legitimization of market-driven innovations which deskill and disenfranchise practitioners. Educational evaluation is not therefore an objective, external, value-free process, but rather is deeply influential in shaping educational philosophy and policy. The conceptual and ethical stance it adopts is influencing the educational debates to a significant degree.

One particular approach the present authors suggest is to move the focus away from external evaluation to one focused on empowering practitioners to self-evaluate. In the case of adult education and training programs the focus of judgment can move from the evaluators to the practitioners and the former can find a new role in supporting the professional development of the latter. Inviting educators to become the key evaluators of educational innovation as opposed to measuring the outcomes in

some external or objective way is of course controversial. For example, it can be argued that educators cannot be objective evaluators as at one level it is their work and effectiveness that is being evaluated. On this account evaluation must be primarily external. However, herein lies the dilemma not just of evaluation but also of perceptions of teaching and learning. On the one hand, there is increasing pressure to reduce teaching to merely implementing a proven or tested program of instruction. On the other hand, the literature of educational improvement has come increasingly to emphasize that the quality of student learning has to be seen in relation to the quality of teachers' learning.

This view of educators' professional learning emphasizes that the quality of teaching is closely bound up with the capacity of teachers to make professional research-based judgments on their own practice and on the programs and methodologies they are being required to implement. This approach is coming to fruition in initiatives such as peer review and peer observation of teaching, best-practice research scholarships, action research, reflective practice, practitioner-led research, and institution-wide development planning. However, for this to truly happen, the values and methodologies of self-evaluation must be inculcated into specific education and training programs.

Conclusion

Learning from and about evaluation often requires us to change our mental models – to rethink our assumptions and beliefs and to develop new understandings about our programs and evaluation processes. This logically should lead on to an organizational learning approach to evaluation. Such an approach to evaluation would be context-sensitive, ongoing, support dialog, reflection, and decision making at department and organization-wide levels, and contain strong commitments to self-evaluation and practitioner empowerment.

Effective evaluation is a significant contributor to quality but does not necessarily guarantee that those in authority will heed the outcomes of evaluation and take necessary corrective action. According to Kells (1992) institutions and programs can be strengthened substantially through effective evaluation and the basis for choices about the future can be soundly established by a combination of internal self-assessment and unbiased, informed peer review. If adult education and training programs are to gain from evaluation procedures and processes then the proper conditions for securing these improvements must be established. A vision of evaluation for the twenty-first century may be one which is made honest, accurate, and useful by engaging in a partnership with practitioners, people, and programs. Evaluators will be held to a higher standard and will be expected to do good through evaluation. In fact evaluation

must move from just generating findings about specific programs to generating knowledge.

To an increasing extent the evaluation model chosen is influenced by the evaluators' own philosophy about evaluation, although other factors such as time, resources, expertise, and availability of staff also strongly influence procedures used. Most program evaluation experts agree that there is no one best model. It is necessary, therefore, for the program evaluator to select a model which matches the requirements of a situation to produce evaluation findings which are most likely to accurately appraise a program's merits, worth, probity, feasibility, safety, significance, and equity.

See also: Cost Analysis in Evaluation Studies; Curriculum Evaluation; Evaluating Education in Three Policy Eras; Program Evaluation.

Bibliography

- Bhola, H. S. (1998). Constructivist capacity building and systemic evaluation of adult basic education and training in South Africa: One discourse, two scripts. *Evaluation* 4(3), 329–350.
- Braverman, M., Engle, G., Arnold, M. E., and Rennekamp, R. A. (eds.) (2008). *Program Evaluation in a Complex Organisational System: Lessons from Cooperative Extension*. San Francisco, CA: Jossey-Bass/American Evaluation Association.
- Jacobs, C. (2000). The evaluation of educational innovation. *Evaluation* 6(3), 261–280.
- Joint Committee on Standards for Educational Evaluation (1994). *The Program Evaluation Standards*. Thousand Oaks, CA: Corwin Press.
- Kells, H. R. (1992). *Self-Regulation in Higher Education. A Multi-National Perspective on Collaborative Systems of Quality Assurance and Control*. London: Jessica Kingsley.
- MacDonald, C. J., Gabriel, M. A., and Cousins, J. B. (2000). Factors influencing adult learning in technology based firms. *Journal of Management Development* 19(3), 220–240.
- McNamara, G. and O'Hara, J. (2004). Trusting the teacher: Evaluating educational innovation. *Evaluation* 10(4), 467–478.
- Rossi, P. H., Freeman, H. E., and Lipsey, M. W. (1999). *Evaluation: A Systematic Approach*, 6th edn. Thousand Oaks, CA: Sage.
- Stufflebeam, D. L. and Shinkfield, A. J. (2007). *Evaluation Theory, Models, and Application*. San Francisco, CA: Wiley.
- Thackwray, B. (1998). *Effective Evaluation of Training and Development in Higher Education*. London: Kogan Page.

Further Reading

- Bezzi, C. (2006). Evaluation pragmatics. *Evaluation* 12(1), 56–76.
- Fetterman, D. M. (2001). The transformation of evaluation into a collaboration: A vision of evaluation in the 21st century. *American Journal of Evaluation* 22(3), 381–385.
- Greenwood, D. J. and Levin, M. (2007). *Introduction to Action Research: Social Research for Social Change*, 2nd edn. London: Sage.
- Hansen, H. F. (2005). Choosing evaluation models: A discussion on evaluation design. *Evaluation* 11(4), 447–462.
- Haug, P. and Schwandt, T. A. (eds.) (2003). *Evaluating Educational Reforms. Scandinavian Perspectives*. Greenwich, CT: Information Age/Corwin Press.
- House, E. R. and Howe, K. R. (2003). Deliberative democratic evaluation. In Kellaghan, T. and Stufflebeam, D. L. (eds.) *International Handbook of Educational Evaluation*, pp 79–102. Norwell, MA: Kluwer.

- Kushner, S. (2000). *Personalising Evaluation*. London: Sage.
- Kushner, S. and Norris, N. (2007). *Advances in Program Evaluation: Volume 10. Dilemmas of Engagement: Evaluation and the New Public Management*. London: Elsevier Science.
- Marion, F. (1988). Phenomenography: A research approach to investigating different understandings of reality. In Sherman, R. and Webb, R. (eds.) *Qualitative Research in Education: Focus and Methods*, pp 143–161. New York: Falmer.
- McNamara, G. and O'Hara, J. (2008). *Trusting Schools and Teachers: Developing Educational Professionalism through Self-Evaluation*. New York: Peter Lang.
- Micari, M., Light, G., Calkins, S., and Streitwieser, B. (2007). Assessment beyond. *Evaluation* 28(4), 458–476.
- Patton, M. Q. (1997). *Utilization-Focused Evaluation: The New Century Text*. Thousand Oaks, CA: Sage.
- Patton, M. Q. (2001). Evaluation, knowledge management, best practices, and high quality lessons learned. *American Journal of Evaluation* 22(3), 329–336.
- Preskill, H. (2008). Evaluation's second act. A spotlight on learning. *American Journal of Evaluation* 29(2), 127–138.
- Shackman, G. (2008). What is program evaluation? A beginners guide. <http://gsociology.icaap.org/methods/evaluationbeginnersguide.pdf> (accessed August 2009).
- Simons, H. (2004). Utilizing evaluation evidence to enhance professional practice. *Evaluation* 10(4), 410–429.
- Stake, R. E. (2004). *Standards Based and Responsive Evaluation*. Thousand Oaks, CA: Sage.
- Torres, R. T. and Preskill, H. (2001). Evaluation and organizational learning: Past, present, and future. *American Journal of Evaluation* 22(3), 387–395.

Relevant Websites

- <http://www.cgu.edu> – Claremont Graduate University, Institute of Organizational and Program Evaluation Research.
- <http://www.eaea.org> – European Association for Education of Adults.
- <http://www.europeanevaluation.org> – European Evaluation Society.
- <http://www.iveta.org> – International Vocational Education and Training Association.
- <http://www.cse.ucla.edu> – National Centre for Research on Evaluation, Standards and Student Testing.
- <http://www.nzqa.govt.nz> – New Zealand Qualifications Authority (NZQA), National Qualifications Framework.
- <http://reflective.bazaar.org> – The Bazaar, Reflective Evaluation.
- <http://www.edpubs.ed.gov> – US Department of Education Publications.
- <http://www.voced.edu.au> – VECED, an Australian Government Funded Operation.