International Review of Research in Open and Distributed Learning

Book Review – Evaluation in Distance Education and E-Learning: The Unfolding Model

Mary F. Kennedy

See table of contents

Explore this journal

April – 2009

Book Review - *Evaluation in Distance Education and E-Learning: The Unfolding Model*

Authors: Ruhe, V. & Zumbo, B. D. (2009). *Evaluation in Distance Education and E-Learning: The Unfolding Model.* New York: The Guilford Press (306 pages). $34.00

Reviewer: Mary F. Kennedy, Program Evaluation Consultants. Contractual Faculty Member, Athabasca University.

In the preface of this book we are informed that there is a substantive link between Messick's four-facet framework for validity in standardized testing and the use of mixed methods to evaluate distance education courses at UBC. The authors make a case, upfront, for a professional approach to evaluation of distance education courses and programs, one that is rooted in the history of program evaluation theory. Certainly, there can be no argument with their stance on this issue.

As we move into the text proper, the authors claim the following as unique features of this text: the need for comprehensive evaluation in light of increasing use of innovative technologies in distance education and e-learning, the explication of a model that can be adapted to local needs, the inclusion of practice tools and strategies for conducting evaluations, and the results of piloting their model in a distance education and an e-learning course.

The authors differentiate between distance education and e-learning. However, today's distance educator, at least in the developed world, lives in one amalgamated world of distance/online learning, which is inclusive of blended learning, Web 1.0, and Web 2.0 technologies. I don't think such a distinction is justified.

The rapid expansion of distance education and the rapid structural, technological, and organizational changes that accompany it are provided as the rationale for increased emphasis on evaluation. The authors conclude Chapter One by presenting their unfolding model, which incorporates scientific evidence and relevance/cost-benefit data within underlying values and unintended consequences.

Chapter Two presents key existing evaluation models. Program evaluation is identified as a transdiscipline. The authors use Messick's framework, along with that of Atkins and Christie, to categorize existing program evaluation models according to their predominant focus on methods, values, or consequences. Most major models are discussed briefly, although Stake's Responsive Model seems to be ignored in favour of his earlier Countenance Model. There is evidence that all

models presented, despite having one major focus, incorporate the other foci as well. However, the authors make a case that Messick's framework is the most comprehensive in that it incorporates methods issues, values issues, and consequences issues equally. They made a statement earlier which has application here: "In our review of evaluation models in distance education, we found almost no models that focus on bringing forward underlying values into the foreground" (p.13). My observation is that Stake's Responsive Model (ignored by the authors) does exactly that.

In Chapter Three, the authors claim to review twelve evaluation models, which they see as specific to distance education, although they stretch their own framework by including Kirkpatrick's training evaluation model, a model with no distance education focus at all. Also, this chapter actually presents seventeen evaluation models, not the twelve indicated earlier by the authors, which leads to some confusion for the reader.

They once again group these seventeen models in accordance with their earlier foci of scientific evidence, values, and consequences. The analysis of each of these models is sound, pointing to both strengths and weaknesses in terms of their application. Table 3.1 rates the seventeen models against Messick's framework, demonstrating that the dimensions of his framework recur in many of the distance education evaluation models.

In Chapter Four, the authors present Messick's validity framework. At the core is the assumption that test validity and program evaluation overlap, and the overlap area consists of Messick's framework. The authors see their unfolding model as "a program evaluation model grounded in the science of test assessment and educational measurement" (p.75). The unified concept of validity incorporates evidential, values, and consequential realms, and construct validity is seen as the unifying force that links the various realms together. The authors conclude that comprehensive program evaluation requires collection and analysis of multiple types and sources of data, from multiple value perspectives.

Chapter Five moves into the action phase of planning the evaluation. The authors make reference to the timing of an evaluation study in relation to the program's stage of development. Inclusion of Owen and Rogers (2007) five evaluation forms (each related to a specific program development stage) would have strengthened this section. They present a series of sections on ethical review processes, which are aimed solely at university faculty research and/or evaluation studies. These sections confirm my opinion that the authors see program evaluation mostly as an academic research pursuit. That view denies the large number of evaluators and evaluation studies that are carried out in all kinds of educational and training environments on a paid, contractual basis. In fact, I would estimate that more program evaluation is done by external, contractual evaluators than within academic communities.

They then present their unfolding model as a road map in developing a workable evaluation plan across five dimensions: scientific evidence, relevance, cost-benefit, values, and unintended consequences. Their overview deals with issues of blending quantitative and qualitative methods

into a mixed method approach, the delineation of potential data sources, data access, sampling, and generalizability versus particularization.

The next two chapters present considerable depth of information on the gathering of scientific evidence and on the search for values and consequences. These chapters provide an overview of data collection methods and analysis, such as is found in any evaluation or research text. Chapter Six deals with scientific evidence, and it provides guidelines for the choice and development or use of many data collection approaches, including surveys, focus groups, interviews, test scores, document analysis, and cost-benefit analysis. Despite the superiority claim of a mixed method approach, I detected, I think, a slight bias in terms of lack of rigor of qualitative data, which is not surprising given the authors' leaning toward quantitative research and indeed their use of Messick's framework. Nonetheless, this stance is irritating to truly qualitative researchers/evaluators.

Chapter Seven deals with value-laden data and with unintended consequences. Values are integrated with distance education benchmarks and standards, and possible data sources are identified. Sample survey/interview questions and data analysis methods are presented in relation to underlying values. Unintended consequences, too, are enumerated, and sample survey/interview questions and data analysis methods are included briefly. Short sections on enhancing validity of findings, reporting results, and making recommendations follow.

Chapter Eight presents findings from the evaluation of two post-secondary distance education courses. Data collection methods for both evaluations were similar, as were data analysis procedures. The summary of evaluation findings in each example demonstrates the integration of scientific evidence with values and unintended consequences. The authors claim these applications of their unfolding model result in measures of merit and worth, despite the diversity of needs of stakeholders and inherent differences in the design and implementation of the courses.

Chapter Nine brings it all together. It reviews yet again Messick's framework and its potential for guiding program evaluation studies, providing a conceptual bridge between assessment and program evaluation. However, as an evaluator, I am not quite sure why there is any need for such a bridge. The authors emphasize the adaptability of their unfolding model to any distance/e-learning course or context.

My summary thoughts? The book is interesting and could serve as a beginners guide to program evaluation. However, it emphasizes course evaluation over program evaluation throughout; and it views evaluation too much from an academic perspective. Furthermore, though their unfolding model is interesting from a theoretical and an application perspective, as an evaluator with many years experience using a great variety of evaluation models, I see no real reason to embrace this model as the primary one for evaluation of distance education and e-learning.

Athabasca University