Section III

Assessment and Evaluation

J. Michael Spector and M. David Merrill

This third section of the *Handbook* is focused on new methods, technologies, and tools to support educational assessment and evaluation. How methods, technologies, and tools have matured and new ones been developed and validated in practical settings are described in several chapters. The intent is to report what is new and avoid repeating what has been published in previous editions of the *Handbook* (Jonassen, 2004; Jonassen, Harris, & Driscoll, 2001; Spector, Merrill, van Merriënboer, & Driscoll, 2008). This section provides an up-to-dateness of research about and with new and emerging assessment and evaluation technologies.

The section begins with a chapter that addresses a topic not previously covered in the Handbook but one that is of global concern—namely, addressing the costs and benefits of educational technology. Thomas Luschei notes the general lack of research about actual cost–benefit analyses of large-scale technology innovations in education and training. The merits and limitations of cost-effectiveness, cost–benefit, cost–utility, and cost–feasibility analyses are described in general. Luschei then describes the methodological issues and challenges of applying these methods in actual settings. More research in this area is encouraged, especially in developing countries where the economic benefits from a cost-multiplier perspective are likely to be significant.

The chapter by Jennifer Hamilton and Jill Feldman on evaluation is also completely new. The traditional perspective on program evaluation is that what is reported is relatively straightforward and simple—namely, the degree to which the program objectives were met. However, there is an increasing interest in being able to connect aspects of an implementation with specific outcomes in order to explain why specific outcomes were or were not obtained. Moreover, the focus has shifted from a simple analysis of outcomes to an ongoing evaluation of a program as it evolves from planning through development and then to implementation and deployment. Formative evaluations along the way have become increasingly significant as evaluators take a more active role in helping program managers achieve intended objectives. Formal methods to evaluate the fidelity of implementation are now required by many funding agencies so that evaluators are positioned to explain why a program did or did not achieve intended objectives. This chapter presents researchers and developers with a practical guide to plan and conduct meaningful program evaluations.

There follow two chapters focused on assessment in domains previously not covered in the *Handbook*—namely, informal learning environments and problem solving. Savenye identifies two areas of informal learning that have received considerable attention in the research literature—museum learning and informal science education. It is well established that much important learning occurs outside school settings. Some informal learning activities can be planned and also assessed. Both quantitative and qualitative methods for such assessments are described by Savenye. Jonassen addresses how to assess problem-solving skills. While research on problem solving has been covered in previous Handbooks, methods to assess improvements in problem solving as a result of training and education have not been addressed. Jonassen provides an overview of assessment methods and coding schemes pertinent to assessing critical cognitive skills, causal reasoning, and other problem-solving skills.

The chapter by Ifenthaler and Pirnay-Dummer on model-based assessment methods and tools is well aligned with the issues raised by Jonassen as it reports on various tools aimed at supporting the kinds of assessments that Jonassen argues are needed for the domain of problem solving. Of particular note is the development of Web-based tools reported in this chapter and the shift that has occurred from a focus on summative assessment to emphasis on formative assessment and dynamic feedback to learners while working on complex and challenging problems.

J.M. Spector, Ph.D. (✉)
University of North Texas, Denton, TX, USA
e-mail: mike.spector@unt.edu

M.D. Merrill
Utah State University, Logan, UT, USA

This sequence of chapters on new areas of concern (informal learning and problem solving) and new assessment tools (Web-based tools that provide dynamic formative feedback) is followed by a reflective chapter on the value and significance of performance assessment. While this topic has been addressed many times and in many contexts, Andrews and Wulfeck argue that performance assessments have returned as the primary means of determining the efficacy of training and how well learners have mastered intended objectives. Many methods and tools to support performance assessment are reviewed as the authors argue that technology now makes it cost-effective to conduct meaningful and predictive performance assessments in many different contexts.

The chapter by Valerie Shute and Yoon Jeon Kim on formative and stealth assessments exemplifies how technology has empowered performance assessment to the extent that nonintrusive formative assessments can be built into a technology-based learning environment. How dynamic formative stealth assessments have been and can be effectively accomplished is described in this chapter.

There follow two chapters that complete this section of the Handbook. One pertains to the task of evaluating information and communication technology (ICT) competencies. Evaluating ICT competencies is important for planning learning environments and for properly preparing teachers, trainers, and students for those environments. However, technologies change quite rapidly and knowledge and skills vary widely from one region to another, so evaluating ICT competencies remains a challenge. Tristan and Ylizaliturri-Salcedo argue that ICT competencies represent a subset of digital literacy competencies, and are essential for the twenty-first century workers. Several approaches for assessing these skills are presented and discussed in this chapter. The final chapter in this section by Kaufman and colleagues then describes how one can and should use data to drive decision making in the classroom. It is clear that commercial enterprises are using data gathered from past purchases and those with similar characteristics to drive how they advertise to specific Internet customers. Likewise, similar technologies and methodologies make it possible to use data to personalize learning and instruction. A data-driven decision-making model is presented to guide such efforts in the future.

The practice of educational technology as a professional discipline requires reliable and effective methods, tools, and technologies to support assessment and evaluation. Professional practitioners need to be able to assess individual learners and learning experiences and evaluate the efficacy of innovative programs. In order to maintain progress, the discipline requires ongoing and cumulative research on assessment and research methods and tools as well as research on specific innovations and implementations. For this reason, we have placed this assessment and evaluation section of the *Handbook* in volume one along with foundations and methods as part of the core discipline of educational technology.

References

Jonassen, D. H. (Ed.). (2004). *Handbook of research on educational communications and technology* (2nd ed.). Mahwah, NJ: Erlbaum.

Jonassen, D. H., Harris, P., & Driscoll, M. P. (Eds.). (2001). *Handbook of research on educational communications and technology* (1st ed.). New York, NY: Taylor & Francis.

* Spector, J. M., Merrill, M. D., van Merriënboer, J. J. G., & Driscoll, M. P. (Eds.). (2008). *Handbook of research on educational communications and technology* (3rd ed.). New York: Routledge.

* An asterisk next to a reference entry throughout this Handbook indicates a reference that the author(s) considers to be central to the topic.