Pinpointing Where to Start: A Reflective Analysis on the Introductory Evaluation Course

Robyn Thomas Pitts
University of Denver

Abstract: *This reflective analysis details four approaches to an introductory course for evaluation learners within a methodologically focused graduate-level program on statistics, measurement, and research design. Evidence of student learning outcomes, or SLOs, was utilized within Gibbs' reflective cycles to redesign the course using Fink's integrated course design process. The purpose of each approach varied along a theory-practice continuum, including theory, theory-to-practice, practice, and evidence building. The purpose, SLOs, and learning experiences of each approach are accompanied by longitudinal reflections on evaluation learners, course purposes, and the creation of a multi-course learning progression. This exploration offers perspectives and lessons learned that may assist new and experienced instructors in determining how an introductory course may best fit the learning needs of their students.*

Keywords: *assessment in higher education, evaluation education, evaluation educator, Gibbs reflective cycle*

Résumé : *La présente analyse réflexive décrit quatre approches possibles pour un cours d'introduction en évaluation dans le cadre d'un programme de cycle supérieur de statistiques, de mesure et de design de recherche, axé sur la méthodologie. Des données liées aux résultats d'apprentissage des étudiantes et étudiants ont été utilisées dans le cadre des cycles réflexifs de Gibbs pour refondre le cours en utilisant le processus de conception de cours intégré de Fink. L'objectif de chaque approche varie en fonction d'un continuum théorie-pratique, y compris la théorie, la théorie à la pratique, la pratique et l'établissement de preuves. L'objectif, les résultats d'apprentissage des étudiantes et étudiants et les expériences d'apprentissage liées à chaque approche sont accompagnés de réflexions longitudinales sur les apprenantes et apprenants en évaluation, les objectifs de cours et la création d'une progression d'apprentissage étalée sur de multiples cours. Cette exploration propose des perspectives et des leçons apprises qui pourraient aider les formatrices et formateurs, tant en début de carrière qu'avec de l'expérience, à déterminer la façon dont un cours d'introduction peut le mieux répondre aux besoins d'apprentissage de leurs étudiantes et étudiants.*

Corresponding author: Robyn Thomas Pitts, Research Methods and Statistics, Morgridge College of Education, University of Denver, 1999 East Evans Avenue, Denver, CO, 80208; Robyn.ThomasPitts@du.edu

Evaluation is a discipline whose theories and practices are applied across myriad settings. Its enigmatic, transdisciplinary nature (Scriven, 2003) provides a considerable impediment to teaching it. When there is no consensus about what constitutes "good" evaluation, educators are challenged to characterize "good" evaluation education. Within the growing body of research (see King & Ayoo, 2020), there are no established purposes, goals, and learning outcomes for graduate study in evaluation; educators disagree about whether theory or practice provides the best entry point into graduate-level training; and learners come from diverse disciplinary backgrounds and prior knowledge of evaluation, ranging from those who lack methodological training to those with years of experience as accidental evaluators. Subgroups have different goals for learning and doing evaluation (Gullickson, 2020), and the number and type of courses needed to prepare future evaluators are unclear. As a result, learner experiences vary considerably across graduate programs (Gullickson et al., 2019).

Educational standards matter because program and course design follow a backward design: Educators start with the end goal in mind. Three other key design principles for high-quality learning include alignment of curriculum, instruction, and assessment; emphasis on learning that results in long-term value; and differentiation for inclusion and accessibility (Fink, 2016). While backward design and alignment focus on understanding course content and pedagogy, the latter two principles focus on customizing learning experiences by understanding who learners are. Integrated course design utilizing these principles advances significant learning, achieved through engaging, high-energy learning processes that result in lasting change and long-term value (Fink, 2013). Since learning occurs across courses, and learners build new knowledge on pre-existing cognitive structures (Gurlitt, 2012), programs of study benefit from careful consideration of cross-course learning through curriculum mapping and learning progressions (Duncan & Hmelo-Silver, 2009; Duschl et al., 2011). Early learning experiences are often instructor-led through guided instruction that introduces new concepts and models ways of thinking about new material. Subsequent learning gradually transfers responsibility for learning from instructor to learners though collaborative learning and independent practice (Fisher & Frey, 2008).

Here I explore four approaches to an introductory evaluation course that vary by focus: theory, theory-to-practice, practice, and evidence building. I used the Gibbs (1998) reflective cycle to make sense of my experiences re-designing the course. I offer hard-earned lessons to help instructors determine how an introductory course may best fit students' learning needs, provide a summary of literature regarding evaluation education and an overview of the Gibbs reflective cycle, and then reflect on the learners, course purpose, and progression of learning across the approaches, concluding with thoughts for university-based evaluation educators.

TEACHING EVALUATION

At the broadest level, evaluation education is facilitated through many forms of professional development (e.g., workshops, conference training, and in-person or online coursework), with university-based programs as the central setting for in-depth, longitudinal learning (Gullickson et al., 2019). Concerted effort has been made to map, update, and compare university-based evaluation training offerings in North America (see Hunter & McDavid, 2018, 2019; LaVelle, 2018, 2019). The International Society for Evaluation Education (ISEE) provides a professional community of practice that convenes through online and in-person meetings to share research findings and create collaborative working groups. In addition to engaging with other educators, instructors can leverage scholarly literature for instructional design.

METHODS

Following an interpretivist approach (Bakker, 2010), I explored four approaches to an introductory evaluation course, critically considering the value and meaning of the learning experiences. This reflective analysis was undertaken through the typical procedures of assessment in higher education, or HED (Banta & Palomba, 2014; Suskie, 2009), utilizing SLO evidence to drive improvement of educational programs (Kuh et al., 2015). In a *New Directions for Evaluation* issue on evaluation in HED, Rickards and Stitt-Bergh (2016) describe SLO assessment as "utilization-focused outcome evaluation with goals that include program/organization improvement and an integration of evaluative thinking in the program and organization" (p. 7).

Positionality

My teaching philosophy is structured around the Gradual Release of Responsibility Model (Pearson & Gallagher, 1983) that promotes learning progressions along a scaffold of cognitive complexity (Dreyfus & Dreyfus, 1980; Fisher & Frey, 2008). I have engaged in SLO assessment at the course and program levels for the past 13 years—as an educator and new teacher trainer, an educational administrator, and an assistant professor and assessment fellow. I have worked as an instructional designer, instructional coach, and educational evaluator in K-12, medical education (MedEd), and HED. As a researcher, I study university-based evaluation training. Though I value structure, I view learning as an idiographic, place-based unfolding that should not be intentionally constrained: Learning experiences should be facilitated in alignment with desired competency development along a general continuum of increasing expertise. My training and experiences in teacher education and MedEd provide a basis for my use of SLO evidence to improve student learning experiences.

Context

Our three-credit-hour introductory evaluation course is taught within a private university on the quarter system, an academic term with 10 class sessions (each

2 hours and 20 minutes in length, generating around 23 hours of seat time and 45–70 potential hours of non-contact study time). Our learners include master's and doctoral degree-seeking graduate students. Though our program is situated within a college of education, our learners are not trained in the substantive knowledge base of education; rather, they study various social science research approaches and tools that can be deployed across a range of applied settings. Our courses are also designed to serve graduate students across the college of education. Prior to 2019, our program offered a single course on evaluation. We are currently developing more robust evaluation training opportunities through multiple evaluation-oriented courses and field experiences.

Data collection and analysis

The data for each reflective cycle included SLO evidence collected through artifact analysis from course-embedded assessments (i.e., intake survey, self-appraisals, artifacts from active learning activities, rubric ratings and narrative comments from presentations and reports, and end-of-course examinations and projects). I also recorded anecdotal notes immediately after each teaching session, engaged in discussions with individual learners, and hosted small-group discussions.

During each course, I analyzed formative assessment data from weekly written reflections or quizzes to identify areas for clarification and remediation. I presented these trends at the outset of each session to guide whole-class discussions of hotspots, tricky concepts, and hard-to-apply skills, discussed these trends and areas for development with my teaching assistants in weekly meetings, and analyzed these evidence sources to generate assessment reports of program-level SLOs. I also discussed my findings with other evaluation educators, benefitting from reviewing and adapting resources they shared. Finally, I attended ISEE meetings and set up conversations with other evaluation educators to discuss detailed issues and address teaching challenges.

Between courses, I examined the learning experience in each approach to the course by following the Gibbs reflective cycle, a structured reflection process for learning from experience that proceeds through six steps: describe the experience without judgment, identify your reactions and feelings during and after the experience, make value judgments about what was good or bad within the experience, make sense of the experience through analysis, draw general and specific conclusions about what else could have been done, and develop an action plan to carry learning forward. To successively approximate the course toward one that would achieve meaningful and long-lasting learning for students, I analyzed evidence of SLOs, integrated new curricular resources to address hotspot areas, and pulled literature on teaching strategies from articles on evaluation education and teaching in HED. I also identified helpful online resources (e.g., didactic presentations on various topics and course syllabi from other programs), using search engines and the library of the teaching of evaluation topical interest group.

FINDINGS AND DISCUSSION

Tables 1–4 illustrate how changing the course purpose altered the SLOs and learning experiences. Each table represents the product of my Gibbs reflective cycle as an action plan for carrying learning forward. Tables include a narrative summary of my longitudinal reflections on learners, course purposes, and the creation of a multi-course learning progression. My goal in redesigning the course using the integrated course-design process (Fink, 2016) was to improve the quality of the student learning experience based on course parameters (i.e., a three-credit-hour course meeting weekly for ten weeks). I also sought to position students' learning about evaluation within the appropriate sequence of their graduate study in order to best achieve our program-level SLOs. The fourth and final approach (Table 4) represents the most appropriate entry point into studying evaluation for learners within our program.

Evaluation learners

Integrated course design around significant learning experiences (Fink, 2013) hinges on understanding who learners are. Over the past three years (2017–2019), the number of students in the introductory evaluation course ranged from 18 to 22 each year, including 36% master's-level methodology students, 52% doctoral-level

Table 1. Approach 1: Introductory evaluation as a survey of evaluation theories

	Survey of evaluation theories (2016)
Purpose	A survey of evaluation theorists and theories
Student learning outcomes (SLOs)	1. Understand what evaluation is and how it differs it from research 2. Understand standards for evaluation and some theories/theorists from the 1960s through today 3. Explicate how program evaluation reflects the evaluator's theoretical orientations
Learning experiences and assessments	Weekly student-led presentations on evaluation theorists • How theorists entered the field • Where and from whom they studied • How their approach connected with other branches of the tree • Main contributions of their work to the field of evaluation Class videoconferences with evaluation theorists Culminating project • Group evaluation plan from a use, methods, or value lens (groups) • Reflection on how theory influenced plan (individuals) • Presentation to peers in class

Table 2. Approach 2: Introductory evaluation as translating theory-to-practice

	Translating theory-to-practice (2017)
Purpose	An introduction to evaluation theory and its role in evaluation design
Student learning outcomes (SLOs)	1. Differentiate evaluation from research 2. Appraise and apply evaluation theories 3. Analyze evaluation reports using knowledge of theory 4. Develop an evaluation study prospectus 5. Apply standards and guiding principles to evaluation scenarios
Learning experiences and assessments	Active learning (e.g., whole- and small-group discussions, evaluating cookie logic and other activities [Preskill & Russ-Eft, 2005]) Personal reflection on what evaluation is and related personal experiences (rubric) Critique of two evaluation reports Culminating project • Annotated bibliography on one evaluation theory • Evaluation study prospectus (with peer review) using theory • Presentation to peers and community partners in class

methodology students, and 12% non-methodology doctoral students who tend to matriculate into leadership positions within educational organizations and non-profits. Among the methodology students, the relative proportion of master's- and doctoral-level students ranged from 25% to 54%, with an overall average of about 40% master's-level students. Methodology learners aspire to work as applied statisticians, measurement specialists, researchers, and/or evaluators. The majority of students taking the course did so to fulfill a curricular requirement within their program of study, with only about 7% (4/59) identifying as evaluators (or aspiring evaluators) at the course outset. This increased to about 22% (10/59) upon completion of the course, a gain of two to four students each year. Of these evaluation-focused graduate students, 70% are master's-level learners. Based on an intake self-appraisal, almost all students indicated minimal to moderate knowledge of evaluation, minimal to moderate experience doing evaluation, and virtually no knowledge of evaluation theory, approaches, or frameworks.

Understanding these frequencies is important, as different student subgroups have different needs for learning about evaluation. Since 78% of the students in the introductory course did not aspire to careers in evaluation, they did not require competency in designing evaluation studies utilizing evaluation theory. Learners of methodology at the master's level needed to develop competencies to prepare them to work on research and evaluation teams. Education doctoral students in non-methodological programs needed an awareness of the field, including how

Table 3. Approach 3: Introductory evaluation as practice via team-based service learning

	Practice via team-based service learning (2018)
Purpose	A practical introduction to the evaluative process
Student learning outcomes (SLOs)	1. Analyze the quality of evaluation studies to identify strengths/weaknesses/unknowns and suggest improvements 2. Design evaluation studies with faculty and community partners using knowledge of diverse modes of inquiry 3. Communicate study design and findings with clarity, both orally and in writing according to APA format and for clients 4. Demonstrate professionalism by employing research ethics when engaging with community partners

Learning experiences and assessments	Active learning (e.g., whole-group discussions, evaluating cookie logic [Preskill & Russ-Eft, 2005], presentations, cross-team communities of practice for discussion and feedback)
	Culminating project
	• Evaluation plan (with peer review) following the CDC self-study guide (CDC, 2011)
	  • Technical report
	  • Stakeholder-friendly report for community partners
	• Peer and self-assessments
	• Presentation to peers and community partners in class

to partner with evaluation specialists. And for the 22% of students desiring careers in evaluation, further coursework and applied experiences were necessary to build on the foundational knowledge of the introductory course. These learner needs suggested that the introductory course should centre on positioning new learning about evaluation relative to prior learning about social science research, particularly by co-constructing areas of similarity and difference to ground learners' perceptions of the scope and utility of evaluation services.

Purpose of the introductory course

The scope of the course, which evolved in response to integrated course design processes (Fink, 2013), was initially theoretical, with a small applied assignment. After the course, students could not readily translate their learning about evaluation theory into practice because they lacked an understanding of working with stakeholders to identify information needs, describe the program context, and generate recommendations for future action. The scope of the second, theory-to-practice approach was too broad: The 10-week course lacked time to engage

Table 4. Approach 4: Introductory evaluation as systematic evidence building

	Systematic evidence building (2019)
Purpose	An introduction to designing evaluation studies A first course within a three-course sequence on evaluation
Student learning outcomes (SLOs)	1. Elaborate ways in which research and evaluation are dis/similar 2. Analyze requests for evaluation services to determine the domain of the evaluation questions and methods 3. Develop an evaluation matrix with a research component or a research matrix with an evaluation component
Learning experiences and assessments	Weekly responses to open-ended questions on key ideas within each domain of evaluation questions and methods, identification of information needs and generation of evaluation matrices for cases aligned with weekly topic/readings, small-group discussions of articles Culminating project • Reflection on careers in evaluation (i.e., analyze job descriptions; identify evaluators working in types of evaluator roles) • Evaluation prospectus with a research component or a research prospectus with an evaluation component • End-of-course examination

deeply and critically with evaluation theory and apply it to study designs. The third, practical approach was challenging to implement because some students struggled to practice evaluation while learning about it for the first time. The fourth approach moved away from issues of evaluation theory and practice altogether: Focusing on the scope and utility of evaluation and its relationship to social science research, it facilitated a more robust introduction to the field as a systematic means of determining the merit, worth, or significance of an evaluand. This arguably narrow focus was possible only because we increased the number of evaluation courses to enable more robust study of evaluation, theory, and practice.

Introducing a learning progression

To address some issues related to the purpose and scope of the course, SLOs initially packed within a single course were distributed across multiple courses in a mapped learning progression. Learning progressions provide coherent environments for learning that unfolds over time by promoting alignment of curriculum, instruction, and assessment (Duncan & Hmelo-Silver, 2009; Duschl et al., 2011); they have been used in K-12 education (Herman, 2013), teacher education (Jin et al., 2015), and medical education (Vantini & Benini, 2008). This year-long, three-quarter learning progression aimed to prepare learners to feel confident working on evaluation teams led by others. Our initial learning progression

proceeded by identifying reasonable prerequisite learning, appropriate SLOs for the introductory course, and a pathway into subsequent courses.

Prerequisite learning

The introductory course was designed to help a heterogeneous student audience position their understanding of evaluation relative to research, a learning goal reasonable only if learners start the course with an operational understanding of social science research. As such, prerequisite courses were added to the sequence so that learners developed an understanding of intermediate statistics (e.g., introductory statistics, correlation, regression, and analysis of variance [ANOVA]), the philosophical assumptions and designs within qualitative inquiry (i.e., narrative, ethnography, case study, grounded theory, phenomenology, and arts-based and criticism/connoisseurship approaches) and quantitative inquiry (i.e., experimental, quasi-experimental, single-subject, correlational, comparative, and descriptive designs), and design/selection of appropriate data collection/analysis methods (e.g., sampling, interviewing, surveying, developing scales, and reviewing ethics).

Student learning outcomes

The scope of the introductory course as systematic evidence building (Table 4) was designed to support learners in making connections between evaluation and social science research. The learning goals further focus on identifying information needs within requests for evaluation services, and developing evaluation matrices (CDC, 2007) that align information needs with domains, research designs, indicators, sources, and methods. Formerly depicted as a pyramid (Rossi et al., 2004), the five domains provide a framework for systematically building evidence by understanding the need for a program and its logic or theory prior to evaluating its processes, outcomes/impacts, or efficiency/optimization. When evidence is missing at a more basal level, findings from evaluative work at subsequent levels may not fully meet the expectations of the standards for program evaluation (Yarbrough et al., 2011). The introductory course increases student understanding of these evaluation-specific aspects of study design and is a launchpad into subsequent learning of evaluation theory and practice.

Subsequent learning

Our program seeks to prepare graduates who can contribute to evaluation teams (master's) and lead studies (doctoral) across numerous applied settings. In the new courses within the evaluation sequence, students worked on teams to conduct an evaluation following Davidson's (2005) approach, bolstered with principles for collaborative and participatory approaches to evaluation (Shulha et al., 2016) as well as culturally responsive approaches (Chouinard & Cram, 2019). Though students completing all three courses were expected to identify the appropriateness of evaluation questions and methods and to facilitate evaluation studies, they were not well prepared to design and implement their own studies. Study design leveraging evaluation theory, engagement with perennial issues of the field, and
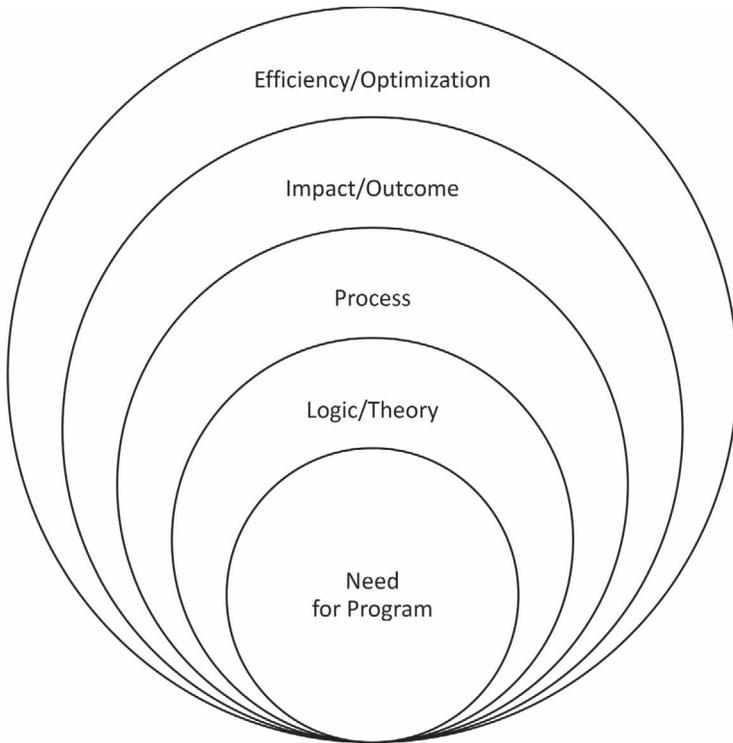
Figure 1. Aggregative evidence building through the five domains of evaluation questions and methods, adapted from Rossi et al. (2018)

research on evaluation were relegated to additional coursework and field experiences beyond the three-course sequence.

CONCLUDING THOUGHTS

Understanding who learners are is key to designing courses that advance meaningful and long-lasting learning. To determine an entry point into learning evaluation within a program of study, instructors can interrogate learner needs and design courses around learning goals that are shared between subgroups. It is unlikely that a single evaluation course can adequately prepare learners to understand the nature of evaluation, help them leverage theory in study design and implementation, and provide them with experiential or service-learning opportunities. Ideas and concepts can be introduced early in a learning sequence, then studied more deeply in subsequent coursework designed to gradually transfer responsibility for learning from instructors to students (Fisher & Frey, 2008). Learning progressions can be helpful tools for generating coherence

in cross-course learning regarding instructional design decisions related to curriculum, instruction, and assessment.

Descriptions of these four approaches to an introductory course only convey their design related to the course purpose, SLOs, and learning experiences; they do not accurately reflect the depth and breadth of learning, which should not be artificially constrained to focus solely on pre-defined competencies. Learning is a sociocultural endeavour rooted in one's personal and cultural biographies (Chouinard et al., 2017), and summaries of course components cannot adequately represent the breadth and malleability of student learning experiences.

While the four approaches are presented as an evolution over time, no single approach is intrinsically superior in its usefulness; indeed, a universal entry point into graduate study of evaluation seems unlikely. The scope of an evaluation course will vary based on program SLOs, the disciplinary domain in which evaluative learning will be applied, and the types of learners for whom the course is designed. This diversity of approaches is simultaneously an asset and a weakness: While learners can select among programs with slightly different foci to accentuate their particular needs and goals, there is no assurance that graduates from across these programs have similar understandings of the field and of how to do evaluation. Entry points into the field will likely always vary; our challenge as evaluation educators is to determine a shared vision for what a degree in evaluation represents and to work backward from that goal in order to engage our learners where they are.

REFERENCES

Bakker, J. L. (2010). Interpretivism. In A. Mills, G. Durepos, & E. Wiebe (Eds.), *Encyclopedia of case study research*. Sage. https://doi.org/10.4135/9781412957397.n180

Banta, T. W., & Palomba, C. A. (2014). *Assessment essentials: Planning, implementing, and improving assessment in higher education* (2nd ed.). Jossey-Bass & Pfeiffer Imprints.

Centers for Disease Control and Prevention (CDC). (2007). *Putting together an evaluation matrix*. US Department of Health and Human Services. https://www.cdc.gov/std/program/pupestd.htm

Centers for Disease Control and Prevention (CDC). (2011). *Introduction to program evaluation for public health programs: A self-study guide*. Centers for Disease Control and Prevention.

Chouinard, J. A., Boyce, A. S., Hicks, J., Jones, J., Long, J., Pitts, R., & Stockdale, M. (2017). Navigating theory and practice through evaluation fieldwork: Experiences of novice evaluation practitioners. *American Journal of Evaluation*, *38*(4), 493–506. https://doi.org/10.1177/1098214016667582

Chouinard, J. A., & Cram, F. (2019). *Culturally responsive approaches to evaluation: Empirical implications for theory and practice*. Sage.

Davidson, E. J. (2005). *Evaluation methodology basics: The nuts and bolts of sound evaluation*. Sage.

Dreyfus, S. E., & Dreyfus, H. L. (1980). *A five-stage model of the mental activities involved in directed skill acquisition*. Defense Technical Information Center. https://doi.org/10.21236/ADA084551

Duncan, R. G., & Hmelo-Silver, C. E. (2009). Learning progressions: Aligning curriculum, instruction, and assessment. *Journal of Research in Science Teaching*, *46*(6), 606–609. https://doi.org/10.1002/tea.20316

Duschl, R., Maeng, S., & Sezen, A. (2011). Learning progressions and teaching sequences: A review and analysis. *Studies in Science Education*, *47*(2), 123–182. https://doi.org/10.1080/03057267.2011.604476

Fink, L. D. (2013). *Creating significant learning experiences: An integrated approach to designing college courses*. John Wiley & Sons.

Fink, L. D. (2016). *Designing courses for significant learning* [PowerPoint Presentation]. Workshop presented at the University of Denver.

Fisher, D., & Frey, N. (2008). *Better learning through structured teaching: A framework for the gradual release of responsibility*. ASCD.

Gibbs, G. (1998). *Learning by doing: A guide to teaching and learning methods*. Further Education Unit, Oxford Polytechnic.

Gullickson, A. M. (2020, June 16). *Snakes and ladders: Evaluation education from a learner's perspective*. International Society for Evaluation Education, Online Conference. https://vimeo.com/429426018

Gullickson, A. M., King, J. A., LaVelle, J. M., & Clinton, J. M. (2019). The current state of evaluator education: A situation analysis and call to action. *Evaluation and Program Planning*, *75*, 20–30. https://doi.org/10.1016/j.evalprogplan.2019.02.012

Gurlitt, J. (2012). Advance organizer. In N. M. Seel (Ed.), *Encyclopedia of the sciences of learning* (pp. 148–151). Springer. https://doi.org/10.1007/978-1-4419-1428-6_157

Herman, J. (2013). *Formative assessment for next generation science standards: A proposed model*. National Center for Research on Evaluation, Standards and Student Testing (CRESST). http://www.ets.org/Media/Research/pdf/herman.pdf

Hunter, M. T., & McDavid, J. C. (2018). *2018 Inventory of Canadian graduate evaluation education*. Consortium of Universities for Evaluation Education. http://evaluationeducation.ca/wp-content/uploads/2018/10/Inventory_Full_Canadian_Graduate_Evaluation_Education_July-2018.pdf

Hunter, M. T., & McDavid, J. C. (2019). Comparison of Canadian and American graduate evaluation education programs. *The Canadian Journal of Program Evaluation*, *34*(2). https://doi.org/10.3138/cjpe.56989

Jin, H., Shin, H., Johnson, M. E., Kim, J., & Anderson, C. W. (2015). Developing learning progression-based teacher knowledge measures. *Journal of Research in Science Teaching*, *52*(9), 1269–1295. https://doi.org/10.1002/tea.21243

King, J. A., & Ayoo, S. (2020). What do we know about evaluator education? A review of peer-reviewed publications (1978–2018). *Evaluation and Program Planning*, *79*, 101785. https://doi.org/10.1016/j.evalprogplan.2020.101785

Kuh, G. D., Ikenberry, S. O., Jankowski, N. A., Cain, T. R., Ewell, P. T., Hutchings, P., & Kinzie, J. (2015). *Using evidence of student learning to improve higher education*. Jossey-Bass.

LaVelle, J. M. (2018). 2018 directory of evaluator education programs in the United States. *University of Minnesota Libraries Publishing, 153*.

LaVelle, J. M. (2019). Educating evaluators 1976–2017: An expanded analysis of university-based evaluation education programs. *American Journal of Evaluation*, *41*(4), 494–509. https://doi.org/10.1177/1098214019860914

Pearson, P. D., & Gallagher, G. (1983). The gradual release of responsibility model of instruction. *Contemporary Educational Psychology*, *8*, 112–123.

Preskill, H., & Russ-Eft, D. F. (2005). *Building evaluation capacity: 72 activities for teaching and training*. Sage.

Rickards, W. H., & Stitt-Bergh, M. (2016). Editors' notes. *New Directions for Evaluation*, *2016*(151), 7–9. https://doi.org/10.1002/ev.20194

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

Rossi, P. H., Lipsey, M. W., & Henry, G. T. (2018). *Evaluation: A systematic approach* (8th ed.). Sage.

Scriven, M. (2003). Evaluation in the new millennium: The transdisciplinary vision. In S. I. Donaldson & M. Scriven (Eds.), *Evaluating social programs and problems: Visions for the new millennium* (pp. 19–41). Lawrence Erlbaum Associates.

Shulha, L. M., Whitmore, E., Cousins, J. B., Gilbert, N., & al Hudib, H. (2016). Introducing evidence-based principles to guide collaborative approaches to evaluation results of an empirical process. *American Journal of Evaluation*, *37*(2), 193–215. https://doi.org/10.1177/1098214015615230

Suskie, L. (2009). *Assessing student learning: A common sense guide* (2nd ed.). Jossey-Bass.

Vantini, I., & Benini, L. (2008). Models of learning, training and progress evaluation of medical students. *Clinica Chimica Acta*, *393*(1), 13–16. https://doi.org/10.1016/j.cca.2008.03.015

Yarbrough, D. B., Shulha, L. M., Hopson, R. K., & Caruthers, F. A. (2011). *The program evaluation standards: A guide for evaluators and evaluation users*. Sage.

AUTHOR INFORMATION

Robyn Thomas Pitts is an assistant professor at the University of Denver. She began her career as an educator and educational evaluator in K-12 and medical education. Currently, she studies evaluation education, teaches graduate coursework on evaluation and mixed methods research, and directs REACH (the Research, Evaluation, & Assessment Collaboratory Hub at DU).