Approaches to Assessment in Higher Education

Colin Madland

2022-03-30

Table of Contents

# Introduction

Assessing learning is both a critical component of the work of teaching in higher education and also a major factor in learners’ experiences of higher education ([Biggs, 1999](#ref-biggsWhatStudentDoes1999); [Woldeab & Brothen, 2019](#ref-woldeab21stCenturyAssessment2019)). Faculty and instructors’ approaches to assessment in higher education are shaped by a variety of factors, including the approaches to assessment that they experienced as learners ([Lipnevich et al., 2020](#ref-lipnevichWhatGradesMean2020); [Massey et al., 2020](#ref-masseyAssessmentLiteracyCollege2020)), pressure from our increasingly technological society to integrate digital tools into the teaching and learning process ([Pellegrino & Quellmalz, 2010](#Xd3220b3c4c74762840488f1ab7714e7e895bc2e)), and the changing needs of 21st-century employers and objectives of higher education institutions, who seek employees and graduates with demonstrated ability in collaboration, creative problem-solving, analytical thinking, and the ability to learn ([Forum, 2020](#Xaf0f34faf6db1662e8da0b9850d3b9cfe5dbeed); [Shute et al., 2010](#ref-shute21stCenturyAssessment2010)), competencies not easily measured in traditional testing formats.

Massey et al., ([2020](#ref-masseyAssessmentLiteracyCollege2020)) contend that instructors in higher education typically have few opportunities to engage in formal preparation for the task of assessing learning, and consequently there is high variability in how instructors assess learning in their courses. This is congruent with Coombs et al.’s ([2020](#ref-coombsSeaSeaCanadian2020)) findings that even preservice teachers in teacher preparation programs are often unprepared for the challenge of assessing learning. If pre-service teachers, who complete a program of formal academic preparation for teaching, are under-prepared, it follows that those who exit doctoral programs with no formal preparation for teaching or assessment ([Lipnevich et al., 2020](#ref-lipnevichWhatGradesMean2020)) will be even less adequately prepared. This lack of formal preparation generally means that higher education instructors assess learning in the only way they know how, which is to follow the example of their supervisors and professors from graduate school.

In her important article, Shepard ([2000](#ref-shepardRoleAssessmentLearning2000)) argues that traditional assessment structures originated in past models of curriculum and instruction which were popular in the early 1900s. These curricular models emphasized the work of psychologists like Thorndike ([Thorndike, 1905](#ref-thorndikeElementsPsychology1905)) and Skinner ([Skinner, 1938](#ref-skinnerBehaviourOrganisms1938)) who viewed the process of learning as being grounded in the mechanistic view of behaviourism where learning is the result of the precise and controlled input of ‘knowledge’ and reinforced with rewards for correct responses. As such, instructors (appropriately) designed their assessments to align with the curricular goals of the time and assessed learning by determining whether or not a learner could provide the single correct response to a given question at a time removed from the instruction. However, in the latter half of the 20th century, when western psychologists discovered the ideas of Lev Vygotsky ([Vygotsky, 1978](#ref-vygotskyMindSociety1978)), curricula began to take a more social-constructivist approach that emphasized higher-order thinking, problem-solving in social contexts, and metacognitive skills over rote memorization.

Unfortunately, it seemed that the efficiencies of testing memory, recognition, and recall through selected-response tests were too deeply embedded in the practices of higher education instructors who resisted changing their assessments to match the new curricular goals ([Shepard, 2000](#ref-shepardRoleAssessmentLearning2000)). Shepard argues for the need to integrate assessment and instruction in such a way as to engage learners in authentic performance tasks more suited to modern understandings of cognition. It appears now that, in the twenty years since Shepard wrote her paper, the goals of early 21st century curricula have continued to diverge from those of the 20th century with the World Economic Forum identifying competencies in collaboration, analytical thinking, creative problem-solving, and the continual learning as being priorities for 21st century employers ([Forum, 2020](#Xaf0f34faf6db1662e8da0b9850d3b9cfe5dbeed)). Models of assessment which mimic and re-inscribe traditional assessment practices and prioritize testing skills in a manner aligned with 20th century curricular models are no longer adequate because they no longer align with the priorities of modern higher education ([Broadfoot, 2016](#X51a879b1b801bbc5ad6e7fc588f4456f2973b2d); [Crooks, 1988](#ref-crooksImpactClassroomEvaluation1988); [Pellegrino & Quellmalz, 2010](#Xd3220b3c4c74762840488f1ab7714e7e895bc2e); [Timmis et al., 2016](#X9783eef1c37e492f1192ce9092108eb526b7f44)).

Furthermore, pressures from our increasingly technological society have also impacted instructors’ approaches to their assessment practice. O’Donnell ([2020](#Xb8ef97dcd64a46920c7df0026dc0b714b2ee92d)) argues that *all* learning in higher education is now mediated in some way by technology, even if it is the superficial use of a word processor an instructor uses to create course materials or a learner uses to compose an essay. Even so, he claims that the use of technology often does little more than increase the efficiency (a term which he declines to define) of existing practices (an example might be reducing the time it takes to score a selected-response exam by using bubble sheets for examinee responses). This critique echoes Timmis et al. ([2016](#X9783eef1c37e492f1192ce9092108eb526b7f44)) who argue that prioritizing superficial characteristics of technology, like ‘efficiency,’ comes at the cost of more innovative applications.

Additionally, the datafication of higher education has made very large data sets available to individual instructors as well as to learning technology administrators. This is often in the form of log data from learning management systems (LMS) which has been used to explore relationships between learner behaviours in the LMS and achievement ([Pardo & Reimann, 2020](#ref-pardoBidirectionalEffectData2020); [Stadler et al., 2020](#ref-stadlerFirstEqualsLog2020)). Shute and Rahimi reported on their exploration of what they call “stealth assessments” ([2021, p. 4](#ref-shuteStealthAssessmentCreativity2021)) where large amounts of data are collected as learners interact in game-based learning environments. They argue that stealth assessments, which learners do not notice because they are woven seamlessly into the learning materials and automatically scored based help to alleviate test anxiety for learners, leading to greater achievement ([Shute et al., 2010](#ref-shute21stCenturyAssessment2010)). Milligan writes about the optimism with which early learning analytics researchers advocated for using big data that ([2020](#X176a232588f05c4bc332011296c94f6bd1d0211))

# A Review of the Literature on Assessment in Technology-mediated Higher Education

## Topic

* Approaches to assessment in technology-mediated higher education

## Problem

* We don’t know how the increased use of technology in higher education has impacted higher education instructors’ approaches to assessment.

## Purpose

* the purpose of this literature review will be to analyze, synthesize and critique the literature since 2010 related to how instructors in higher education approach classroom assessment in increasingly technology-mediated environments.

## Questions

1. What are the major themes or patterns in the literature related to assessment in higher education?
2. What are the major themes or patterns in the literature related to the impact of technology on assessment in higher education?
3. What gaps exist in the literature related to assessment in technology-mediated higher education?

## Defining Assessment

### Historical Definitions

* Among the more influential publications related to modern views of assessment (then usually called “evaluation”) was Scriven’s ([1967](#ref-scrivenMethodologyEvaluation1967)) article in which he drew distinctions between “formative” and “summative” evaluation.
* This distinction was quickly incorporated into Bloom’s ([1968](#ref-bloomLearningMasteryInstruction1968)) ideas related to mastery learning and began to be promoted as a model for educational reform.
* However, by the late 1990s, when Black and Wiliam ([1998](#ref-blackAssessmentClassroomLearning1998)) published their thorough review of the literature, the idea of formative assessment was still not well-defined or implemented. Black and Wiliam framed formative assessment as “encompassing all those activities undertaken by teachers, and/or by their students, which provide information to be used as feedback to modify the teaching and learning activities in which they are engaged” ([1998, pp. 7–8](#ref-blackAssessmentClassroomLearning1998)).
* The National Research Council’s (NRC) 2001 report *Knowing what students know*, advanced understanding of assessment with their definition of assessment as “a process of drawing reasonable inferences about what students know on the basis of evidence derived from observations of what they say, do, or make in selected situations” ([Pellegrino et al., 2001, p. 112](#ref-pellegrinoKnowingWhatStudents2001)) or, more simply, “reasoning from evidence” ([Pellegrino et al., 2001, p. 43](#ref-pellegrinoKnowingWhatStudents2001)), based on Mislevy’s assertion that “test theory is machinery for reasoning from students’ behavior to conjectures about their competence, as framed in a particular conception of competence.” ([1994, p. 4](#ref-mislevyTestTheoryReconcieved1994)).
* The NRC models assessment as a triangle composed of three interdependent components of any assessment (Figure 1): *cognition*, or a model of the domain to be learned; *observation*, or the performance task learners will complete to demonstrate their competence; and an *inference* or *interpretation* of the data produced by the observation. The interdependent nature of the three components requires that both the observation and interpretation components be grounded in the nature of the cognitive model of the domain.

### Modern Conceptions of Assessment

* Since ~2010, there has been a shift in how researchers view assessment towards assessment being a complex, multi-faceted construct in which instructors bring with them a host of experiences and beliefs about assessment, some of which were passed down from their own instructors ([Lipnevich et al., 2020](#ref-lipnevichWhatGradesMean2020); [Massey et al., 2020](#ref-masseyAssessmentLiteracyCollege2020)), but others are based in course and institutional policies, and the social dynamics within their department ([G. T. L. Brown et al., 2011](#X7545160f175b9eeeaf92f7a1f4f6c5a1f4f7d20)).
* Earl ([2013](#ref-earlAssessmentLearningUsing2013)) further clarified the role assessment can play in learning by highlighting a distinction between assessment *of* learning (summative assessment) and assessment *for* learning (formative assessment by way of feedback) and also distinguishing both of those from assessment *as* learning (a subset of assessment *for* learning in which learners employ metacognitive skills to regulate their own learning tasks).
* DeLuca et al. ([2013](#X32b95a94fb3076f8d0c577bd9276160db2e95e5)) argue that there are categories of conceptions exhibited by K-12 preservice teachers: assessment as testing, assessment as format, assessment purpose, and assessment as process. These conceptions are seen as increasingly complex, with those who see assessment as testing believing that assessment is primarily concerned with summative assessment of learning, usually using teacher-created selected-response tests. Those who see assessment as format tend to focus on whether the assessment is a “performance, product, or objectively-scored assessment” (p. 110). Assessment as purpose is delineated according to the summative/formative binary or Earl’s ([2013](#ref-earlAssessmentLearningUsing2013)) assessment *of/for/as* learning model. DeLuca et al. also identify other purposes of assessment such as accountability, gatekeeping, and teacher evaluation. Lastly, assessment as process, which is based on the National Research Council’s description of assessment being a process of reasoning from evidence ([2001](#ref-pellegrinoKnowingWhatStudents2001)).
* Fletcher et al. ([2012](#Xa55a5d631b402fa2c3fda6e3997bc5d48ccd027)) used Brown’s ([2017](#X398061bedf4ca6ea011494eb395f9026963cd42)) abridged *Conceptions of Assessment* (CoA) questionnaire to measure learners’ and instructors’ conceptions as follows: “assessment makes institutions accountable, assessment makes students accountable, assessment describes improvements in student abilities, assessment improves student learning, assessment improves teaching, assessment is valid, assessment is irrelevant and bad, assessment is irrelevant and ignored, and assessment is irrelevant and inaccurate” (p. 122). They report that instructors were more likely than learners to view assessment as consistent and trustworthy methods to understand and improve learning and that learners were more likely to have negative views of assessment and viewed it as a measure of student and institutional accountability.
* Massey et al. ([2020](#ref-masseyAssessmentLiteracyCollege2020)) used DeLuca et al.’s ([2013](#X32b95a94fb3076f8d0c577bd9276160db2e95e5)) framework of conceptions in their study of HE instructors’ conceptions of assessment before and after an instructional development course focussed on assessment. They also considered the idea that there are two general orientations towards assessment in HE, an “assessment culture” and a “testing culture” [Massey et al.](#ref-masseyAssessmentLiteracyCollege2020) ([2020](#ref-masseyAssessmentLiteracyCollege2020)). They report that they saw significant shifts in participants’ conceptions of assessment from more simplistic views of assessment as testing pre-treatment, to more complex and nuanced views or assessment as process post-treatment.

## Assessment Literacy

## Approaches to Assessment

## Assessment and Measurement

### Validity

### Reliability

### Fairness

## Assessment in Higher Education

## Impact of Technology on Assessment in Higher Education

Biggs, J. (1999). What the Student Does: Teaching for enhanced learning. *Higher Education Research & Development*, *18*(1), 57–75. <https://doi.org/drgphk>

Birenbaum, M. (1996). Assessment 2000: Towards a pluralistic approach to assessment. In M. Birenbaum & F. J. R. C. Dochy (Eds.), *Alternatives in assessment of achievements, learning processes and prior knowledge* (Vol. 42, pp. 3–29). Kluwer Academic/Plenum Publishers.

Birenbaum, M. (2003). New insights into learning and teaching and their implications for assessment. In M. Segers, F. Dochy, & E. Cascallar (Eds.), *Opti- mising new modes of assessment: In search of qualities and standards* (pp. 13–36). Kluwer Academic Publishers.

Black, P., & Wiliam, D. (1998). Assessment and Classroom Learning. *Assessment in Education: Principles, Policy & Practice*, *5*(1), 7–74. <https://doi.org/fpnss4>

Bloom, B. (1968). Learning for Mastery. Instruction and Curriculum. Regional Education Laboratory for the Carolinas and Virginia, Topical Papers and Reprints, Number 1. *Evaluation Comment*, *1*(2), 12.

Broadfoot, P. (2016). Assessment for Twenty-First-Century Learning: The Challenges Ahead. In M. J. Spector, B. B. Lockee, & M. D. Childress (Eds.), *Learning, Design, and Technology* (pp. 1–23). Springer International Publishing. <https://doi.org/10.1007/978-3-319-17727-4_64-1>

Brown, G. (2017). *Teachers Conceptions of Assessment - Secondary Schools Long and Abridged*. <https://doi.org/gj4tz6>

Brown, G. T. L., Lake, R., & Matters, G. (2011). Queensland teachers’ conceptions of assessment: The impact of policy priorities on teacher attitudes. *Teaching and Teacher Education*, *27*(1), 210–220. <https://doi.org/c3k8f5>

Coombs, A., Ge, J., & DeLuca, C. (2020). From sea to sea: The Canadian landscape of assessment education. *Educational Research*, 1–17. <https://doi.org/gh5k4z>

Crooks, T. J. (1988). The Impact of Classroom Evaluation Practices on Students. *Review of Educational Research*, *58*(4), 438–481. <https://doi.org/dvd8nf>

DeLuca, C., Chavez, T., & Cao, C. (2013). Establishing a foundation for valid teacher judgement on student learning: The role of pre-service assessment education. *Assessment in Education: Principles, Policy & Practice*, *20*(1), 107–126. <https://doi.org/gj5v98>

Earl, L. M. (2013). *Assessment as learning: Using classroom assessment to maximize student learning* (Second edition). Corwin Press.

Fletcher, R. B., Meyer, L. H., Anderson, H., Johnston, P., & Rees, M. (2012). Faculty and Students Conceptions of Assessment in Higher Education. *Higher Education*, *64*(1), 119–133. <https://doi.org/ctccpq>

Forum, W. E. (2020). *The Future of Jobs Report 2020*. World Economic Forum.

Lipnevich, A. A., Guskey, T. R., Murano, D. M., & Smith, J. K. (2020). What do grades mean? Variation in grading criteria in American college and university courses. *Assessment in Education: Principles, Policy & Practice*, *27*(5), 480–500. <https://doi.org/ghjw3k>

Massey, K. D., DeLuca, C., & LaPointe-McEwan, D. (2020). Assessment Literacy in College Teaching: Empirical Evidence on the Role and Effectiveness of a Faculty Training Course. *To Improve the Academy*, *39*(1). <https://doi.org/gj5ngz>

Milligan, S. (2020). Standards for Developing Assessments of Learning Using Process Data. In M. Bearman, P. Dawson, R. Ajjawi, J. Tai, & D. Boud (Eds.), *Re-imagining University Assessment in a Digital World* (Vol. 7, pp. 179–192). Springer International Publishing. <https://doi.org/10.1007/978-3-030-41956-1_13>

Mislevy, R. J. (1994). Test theory reconcieved. *ETS Research Report Series*, *1994*(1), i–38. <https://doi.org/gjm236>

O’Donnell, M. (2020). Assessment as and of Digital Practice: Building Productive Digital Literacies. In M. Bearman, P. Dawson, R. Ajjawi, J. Tai, & D. Boud (Eds.), *Re-imagining University Assessment in a Digital World* (Vol. 7, pp. 111–125). Springer International Publishing. <https://doi.org/10.1007/978-3-030-41956-1_9>

Pardo, A., & Reimann, P. (2020). The Bi-directional Effect Between Data and Assessments in the Digital Age. In M. Bearman, P. Dawson, R. Ajjawi, J. Tai, & D. Boud (Eds.), *Re-imagining University Assessment in a Digital World* (Vol. 7, pp. 165–178). Springer International Publishing. <https://doi.org/10.1007/978-3-030-41956-1_12>

Pellegrino, J. W., Chudowsky, N., & Glaser, R. (2001). *Knowing What Students Know: The Science and Design of Educational Assessment*. National Academies Press. <https://doi.org/10.17226/10019>

Pellegrino, J. W., & Quellmalz, E. S. (2010). Perspectives on the Integration of Technology and Assessment. *Journal of Research on Technology in Education*, *43*(2), 119–134. <https://doi.org/ggfh8z>

Scriven, M. (1967). The methodology of evaluation. In B. O. Smith (Ed.), *Perspectives of curriculum evaluation*. Rand McNally.

Shepard, L. A. (2000). The Role of Assessment in a Learning Culture. *Educational Researcher*, *29*(7), 4–14. <https://doi.org/cw9jwc>

Shute, V. J., Dennen, V. P., Kim, Y., Donmez, O., & Wang, C. (2010). *21st century assessment to promote 21st century learning: The benefits of blinking*.

Shute, V. J., & Rahimi, S. (2021). Stealth assessment of creativity in a physics video game. *Computers in Human Behavior*, *116*, 106647. <https://doi.org/10.1016/j.chb.2020.106647>

Skinner, B. (1938). *The behaviour of organisms*. Appleton-Century-Crofts.

Stadler, M., Hofer, S., & Greiff, S. (2020). First among equals: Log data indicates ability differences despite equal scores. *Computers in Human Behavior*, *111*, 106442. <https://doi.org/10.1016/j.chb.2020.106442>

Thorndike, E. L. (1905). *The Elements of psychology*. A.G. Seiler.

Timmis, S., Broadfoot, P., Sutherland, R., & Oldfield, A. (2016). Rethinking assessment in a digital age: Opportunities, challenges and risks. *British Educational Research Journal*, *42*(3), 454–476. <https://doi.org/gftz95>

Vygotsky, L. S. (1978). *Mind in society* (M. Cole, V. John-Steiner, S. Scribner, & E. Souberman, Eds.; A. R. Luria, Trans.). Harvard University Press.

Woldeab, D., & Brothen, T. (2019). 21st Century assessment: Online proctoring, test anxiety, and student performance. *International Journal of E-Learning & Distance Education*, *34*(1).