2023-10-10

Epistemologies Of Assessment Instruments

Raenita FENNER
Loyola University Maryland, rafenner@loyola.edu

Peggy O'NEILL
Loyola University Maryland, Poneill1@loyola.edu

Kerrie DOUGLAS
Purdue University, douglask@purdue.edu

See next page for additional authors

Recommended Citation

Fenner, R., O'Neill, P., Douglas, K., & Douglas, E. P. (2023). Epistemologies Of Assessment Instruments. European Society for Engineering Education (SEFI). https://doi.org/10.21427/B8KH-Z263

Authors

Raenita FENNER, Peggy O'NEILL, Kerrie DOUGLAS, and Elliot P. DOUGLAS

EPISTEMOLOGIES OF ASSESSMENT INSTRUMENTS

R. Fenner
Loyola University Maryland
Baltimore, MD, USA
ORCID 0000-0002-2289-870X

P. O'Neill
Loyola University Maryland
Baltimore, MD, USA
ORCID 0000-0002-1026-3483

E. P. Douglas[1]
University of Florida
Gainesville, FL, USA
ORCID 0000-0001-6582-9758

K. Douglas
Purdue University
West Lafayette, IN, USA
ORCID 0000-0002-2693-5272

Conference Key Areas: *Fostering Engineering Education Research, Other Topics in Engineering Education*
Keywords: *epistemology, assessment, quantitative literacy*

[1] Corresponding Author

E. P. Douglas

edouglas@ufl.edu

ABSTRACT

Understanding our epistemological perspective when conducting engineering education research is important for situating the knowledge claims we are making. Depending on that perspective, we may situate the knowledge claims as definitive, representing an absolute Truth, or as contingent, representing a contextualized truth. Traditionally, quantitative research has been identified as positivist, while qualitative research is diverse in its epistemological assumptions, ranging from positivist to interpretivist to Critical and the "posts." Thus, results from quantitative studies are often treated as generalizable, absolute, and decontextualized, while quantitative studies are treated as particular, contingent, and contextualized.

Assessment instruments, being quantitative, are associated with positivist forms of knowledge. We argue that it is more appropriate to treat quantitative assessments as interpretivist. Development of assessments is based on particularized knowledge that is created through a dialogue between the developers and the pilot participants. Interpretation of assessment results is dependent on the particular contexts in which they are used.

In this paper we describe the interpretivist roots of assessment using the example of our current project on developing an instrument for engineering quantitative literacy. In the first phase of this project we have used qualitative content analysis to identify the ways in which quantitative literacy is assessed in first-year engineering courses in the United States. This analysis is contextualized by the particulars of these courses, and the results are contingent on the interpretations we make as researchers. We discuss how this interpretivist perspective carries through the entire project as we create and implement a measure of quantitative literacy for engineering students.

1 INTRODUCTION

At its core, research is the practice of making knowledge claims from empirical evidence. How we make these knowledge claims depends on our views of how knowledge is defined, or our epistemological beliefs. Epistemology is the nature of knowledge; what counts as knowledge, how we understand what knowledge is, and where knowledge comes from (Crotty 1998). Being explicit about one's epistemological assumptions provides transparency in the research process and helps researchers select appropriate methodologies that are coherent with those assumptions (Koro-Ljungberg et al. 2009, Douglas, Koro-Ljungberg, and Borrego 2010). Studying the same phenomenon from different epistemological perspectives can provide a more holistic view that would not be possible from a single perspective (Baillie and Douglas 2014, Douglas, Koro-Ljungberg, and Borrego 2010).

In this paper, our goal is not to report results from our study but to expand notions of epistemological diversity. In particular, we seek to disrupt the assumption of quantitative research as inherently positivist. Modern Western notions of science default to a positivist epistemology, with its assumptions of objective Truth. Quantitative research is viewed as objective and neutral, allowing us to identify that Truth. We argue that in social sciences generally, and in engineering education in particular, even quantitative research should be seen as non-positivist. As an example case, in this paper we discuss how we have considered epistemology in our current research project to develop a quantitative instrument for engineering quantitative literacy.

2 LITERATURE REVIEW

2.1 Epistemology

Broadly speaking, epistemologies can be grouped into four categories: positivist, interpretivist, Critical, and the 'posts' (post-structural and post-modern). Positivism is generally associated with quantitative research, although qualitative research can also be conducted from a positivist perspective. Positivism assumes that there is a single objective Truth that exists independently of humans. The goal of research is to identify that Truth, although empirical claims are always subject to falsification. Positivism has important implications for how research is conducted. In quantitative research, instruments should be reliable to ensure that they consistently measure the construct of interest. Positivism also implies the need for large sample sizes in order to separate the true effect from noise. And statistical analysis is needed to determine how likely it is that our data represents a true effect (e.g., the use of p-values in hypothesis testing). In qualitative research, positivism results in the need to have multiple coders and to calculate the inter-coder reliability. Since there is an objective Truth in the data, two different coders should code the data the same way.

Interpretivism is generally associated with qualitative research, although the thesis of this paper is that quantitative research can also be interpretivist. Interpretivism recognizes that there is no single Truth, but rather multiple truths that are created through people's interaction with the world. Thus, truth is contingent and contextual. Interpretivism in the context of assessment is discussed further in the next section.

A Critical epistemology takes knowledge to not only be contingent, but subject to politics (i.e., the struggle for resources). Thus, Critical research examines issue of power and seeks to disrupt existing power relationships. There are various Critical approaches, each with its own set of tenets. QuantCrit specifically addresses the issue of quantitative data with tenets that neither numbers or categories are value-neutral (Suzuki, Morris, and Johnson 2021). It recognizes that presenting quantitative research as 'objective' can mask the ways it is used to maintain racial hierarchies. 'Post' epistemologies take knowledge to be embedded in grand narratives that describe the way the world 'should' be. The role of 'post' research is to disrupt those narratives.

2.2 The Epistemology of Educational Assessment

Epistemology is not often discussed in engineering education research articles, particularly those with methods using educational assessments or surveys. While it is often unacknowledged, the researchers' epistemology can be understood from how the study's methodology is approached and results discussed (Koro-Ljungberg and Douglas 2008, Koro-Ljungberg, et al. 2009). Are the researchers approaching the study from the perspective that their own interpretation is part of the research process? Or perhaps the researchers make statements that indicate the results somehow speak for themselves and there is no researcher finding. In the case of educational assessment, the researchers' epistemology may be unclear or even seemingly conflicted between the use of an assessment and the conclusions drawn.

Researchers in the measurement community have long communicated that subjectivity and interpretation are inherent in all aspects of educational assessment (e.g., Thorndike and Hagen 1977, Messick 1998), but rarely does one find practical discussion of what these aspects mean for research and educational use (e.g, Smith 1989). Thus, scores resulting from assessment instruments are too often interpreted simply as a valid measure and there is no variability of interpretation (Douglas and Purzer 2016). Many students have been evaluated based purely on a resulting assessment score, without contextualization of the subjectivity of the assessment or the values of the assessment developer. Such an approach is misaligned with the constructivist nature of educational assessment. Hence, discussion of epistemology is warranted to remind engineering education researchers, administrators, and educators to treat assessment scores with a degree of humility and to challenge the community to develop assessments with interpretivist or critical considerations explicitly in mind.

The beginning of any assessment is to define what is to be measured and the scope. Cronbach and Meehl (1955) defined the term *construct* as, “some postulated attribute of people, assumed to be reflected in test performance” (p. 4). Considering no one can see inside another’s mind to know what is understood, thought, or felt, the assessment developer must first decide what the construct definition is, and then how it will be measured. Put another way, educational assessment seeks to define what should be measured, then create opportunities where the student can demonstrate their knowledge, attitude, etc. in a way that assessment users can then use to draw reasonable inferences about what the students know (Pellegrino 2013). In short, the assessment developers construct what will be assessed, how it will be assessed and how it will be used and interpreted. Thus, subjectivity is an inherent characteristic of all assessment, from development through to decisions and resulting consequences. “The test score is not equated with the construct it attempts to tap, nor is it considered to define the construct, as in strict operationism. …the measure is viewed as just one of an extensible set of indicators of the construct” (Messick 1989, p.7). Years later, Messick (1998) wrote more directly, “constructivism is central to the whole enterprise of construct validity” (p.35). While the scientific or mathematical knowledge represented in educational learning goals may have concrete truths (e.g., mathematically, 4+4 equals 8), the process of measuring that mathematical principle is constructed.

The constructivist nature of assessment necessitates that there is no such thing as a perfectly valid assessment (Songer and Ruiz-Primo 2012). Unlike measures used in fields of engineering that reflect a scientific principle (e.g., volume is measured by the same equation around the world), educational assessments attempt to measure constructs that the reality of existence can scarcely be proven. People construct the name, what it means, and how to measure it. Validation is intended to be the evaluative process regarding what inferences can reasonably be made and uses for the scores are justifiable. As we cannot see directly into someone’s mind, there’s no one ‘right’ way to measure what they know and can do. Thus, some scholars have argued that educational assessment attempts to measure something far more difficult than the physical realm measured by engineers (Douglas and Purzer 2016, Wankat et al. 2002). It is precisely the difficulty of measuring something that is not observable that formed the field of educational measurement and psychometric techniques. However, those techniques do not change the inherent nature of educational assessment – from start to finish, the developers and users make subjective decisions. Psychometric methods are the statistics concerned with modelling measurement error. This acknowledgement in the statistical methods that all educational measures are imperfect tools is misaligned when assessment scores are summed up, graded right/wrong and then used from a positivist perspective.

Despite common understanding that all educational assessments are value-laden and subjective, judgments made about students or groups of students based on assessment scores seem to reflect a 'truth' that the score is in fact all the researcher, educator, or administrator need to know (Smith 1989). Mislevy (1994) discussed how researchers of different paradigms (behavioural, information processing, and constructivist) might approach validation studies in terms of evidence collected and inferences sought. Simply acknowledging that researchers have different epistemological understandings is exactly the reason why researchers approach their work from the perspective that they themselves cannot be separated from the research. What evidence is collected and the scientific argument for why that evidence is sufficient is subjective. What constitutes 'enough' or 'good enough' is highly context dependent, as standards argue that the more a test is used for decisions of personal consequence, the more evidence is warranted (AERA, APA, and NCME 2014). Assessment frameworks and standards are intended to increase the principled nature of developing, validating, and using assessments (e.g., Evidence-Centered Design, Mislevy 1994, Argument-Based Approach, Kane 2016).

Ethical approaches to assessment acknowledge the diverse ways people know, experience, and demonstrate their knowledge. From cognitive science research, we understand that learning is mediated by culture, language, and other tools (National Academies of Sciences, Engineering, and Medicine 2018). The sociocultural-informed Evidence-Centered Design approach (Oliveri, Nastal, Slomp 2020) is an assessment design and validation model that explicitly considers how diverse groups of students would experience, understand, and demonstrate their understanding in diverse ways. Without this explicit acknowledgement, the assessment developer runs the risk of assessing students in the way they themselves demonstrate understanding and then holding students from different socio-cultural backgrounds to the same way of knowing.

3 THE CASE – ENGINEERING QUANTITATIVE LITERACY

As the goal of this paper is to discuss the epistemology of assessment, here we provide only a brief overview of our current project to develop an instrument for engineering quantitative literacy. More details are provided in a recent paper (Fenner et al. 2023).

Quantitative literacy (QL) is the ability to engage in context-specific quantitative activities for problem solving. The chapters in the book edited by Gillman (2006) provide an overview of how QL is taught across the curriculum at a number of institutions. There are a number of definitions of QL (AACU 2014, Mayes, et al, 2013, OECD 2012, Sons, 1996, Vacher, 2014, Wilkins, 2010). There is a consensus among these definitions that QL consists of mathematical skills, communication of quantitative information, interpretation and reasoning, and ability to apply these elements in particular contexts. Wilkins (2010) identified three components of QL: disposition, beliefs, and cognition. The cognition component can be further divided into content, reasoning, and communication (Roohr, Graf, and Liu 2014, Kosko and Wilkins 2011), and thus the cognition component encompasses the elements of the other definitions. While a number of instruments have been developed to measure QL skills (see, for example, ETS 2021, Kosko and Wilkins 2011, Zahner et al. 2021), existing instruments were designed for grade school, general college students, or the general adult population.

While there is broad agreement that quantitative skills are critical for engineers, there has been almost no work done on the QL skills of engineering students. Prince and Simpson (2016) have written the only paper which focused entirely on engineering students. The goal for our project is to develop an instrument for engineering QL, which can then be used to assess the QL skills of engineering students.

Our overall project uses the Evidence-Centered Design process (Mislevy, Almond, and Lukas 2003), a comprehensive framework for designing and validating assessments. In this framework one articulates assessment arguments in each of three models in an assessment: Student Model, Evidence Model, and Task Model (Riconscente, Mislevy, and Corrigan 2016). We are currently developing the Student Model, which defines what we will measure, i.e., the variables related to the knowledge, skills, and abilities we want students to learn.

To develop the Student Model, we first developed a definition of QL. Based on our review of the literature, we defined QL as

> *The ability to engage in context-specific quantitative activities for problem-solving and communication by collecting, understanding, processing, interpreting, synthesizing, and displaying numerical information. This definition includes numerical skills and dispositions, and beliefs in quantitative activities.*

We then collected course syllabi, assignments, and exams for first-year engineering courses at five institutions in the U.S. We developed a coding frame based on our definition of QL and the components of QL described above. The course materials were then coded by the first two authors, with the third and fourth authors reviewing the coding as a check on quality. It is important for the discussion below to note that we did not calculate an inter-coder reliability. Codes were discussed among the authors and discrepancies were resolved through consensus.

Out of 125 QL tasks in the course materials, all fell into the Cognitive dimension, with none in Beliefs or Dispositions. Within Cognition, 85.6% were coded as Reasoning, 9.6% as Communication, and 4.8% as Content. Examples of the tasks and further discussion of these results is given in our prior publication (Fenner et al. 2023). We are continuing to develop the Student Model by categorizing the types of QL tasks that are present in our data.

4 EPISTEMOLOGY IN PRACTICE

Initially we implicitly conceptualized this project as positivist. Our definition of QL gave us a specific, detailed description of what constituted QL tasks that we intended to use to characterize tasks in the course materials. Our assumption was that the definition provided a clear, unambiguous way to identify QL tasks that would be valid regardless of the specific document we were analyzing.

However, as we began analyzing the data and discussing the coding, we recognized and questioned our positivist assumptions. Each course was taught in a specific context, to a particular group of students, and thus each course had a particular focus. Our analysis of what constitutes QL in engineering is thus shaped by those contexts. We also realized that, despite its apparent specificity and clarity, our definition ultimately required us to rely on our own interpretive skills in order to assign each task to an appropriate QL dimension. The contextual and interpretive nature of the data is apparent at three levels: course, instructor, and task.

At the course level we noted that most of the content focuses on circuits. While we do not yet know if this focus will be present as we collect materials from more schools, it does impact how QL is ultimately defined. QL skills based on circuits may be different from those based on structures, for example. The instrument we ultimately develop must therefore be understood as contextualized based on how it was developed, and not represent an absolute, objective measure of QL.

At the instructor level, one of the coders teaches courses in linear circuits. Thus, during coding with tasks associated with linear circuits, she became aware of the need to 'bracket' her assumptions about what she would expect students to do, and only code based on what the assignment actually said. This need to bracket one's own biases is an element of qualitative, interpretive research. Bracketing results from the lack of a single, objective Truth that is present in all contexts.

At the task level, we found that our QL definition was not unambiguous and required us to interpret what it meant for a given task. For example, activities related to Kirchhoff's Current Law appeared frequently in the student materials. Kirchhoff's Law can be summarized as: for any node in an electrical circuit, the sum of currents flowing into that node is equal to the sum of currents flowing out of that node. Initially, we assumed that any task or assignment involving Kirchhoff's Law was a QL task because Kirchhoff's Law requires some sort of mathematical understanding as marked by the terms "equal to" and "sum of." However, the interpretation of quantitative literacy as an interpretivist construct led us to question the assumption that all tasks requiring mathematical skills are necessarily QL tasks. Through our encounters with different types of tasks, we came to the realization that some activities may require students to apply mathematical concepts or solve problems, while others may demand a different set of skills altogether. For instance, while one task might call for students to work through a problem using Kirchhoff's Law, another might ask them to explain this law in simpler terms suitable for a non-expert audience such as grandparents or readers of a newspaper article. Thus, whether or not a particular tasks involves QL requires an interpretation of what QL means in that context.

5 SUMMARY AND ACKNOWLEDGMENTS

Our experience shows how quantitative research in engineering education is interpretive. While most of engineering is positivist, e.g., a given circuit has predictable behavior, the same cannot be said of research involving human subjects. Our experience with QL assessment development shows the ways in which quantitative instruments are more appropriately considered from an interpretivist perspective. Thinking of QL from a positivist perspective implies a singular definition, with assessment results that objectively identify a student's QL skills. However, our literature review and experience with this project show that assessment results need to be understood as contextualized and subjective (i.e., subject-focused). As we proceed in our project, our assessment of QL skills will need to be considered as situated. They will be based on a particular way we have defined QL and operationalized that definition. Other definitions and operationalizations, equally relevant, could have different results. In addition, we will need to interpret the results in light of the students who respond to the instrument.

Considering assessment as interpretivist rather than positivist has important implications. Traditional measures of quality, such as reliability, are not sufficient for ensuring appropriate use of assessment instruments. Instead, researchers need to be reflexive about the context in which the assessment is occurring and their own biases. In selecting instruments, researchers should ask themselves questions such as: In what context was the assessment instrument created? How does that compare to the context in which I am using the instrument? How do I understand the topic and how has that influenced which assessments measures I chose? In interpreting the assessment results, questions include: What contextual factors might affect these results? How does my understanding of the topic influence my interpretation of the results? How would students from different socio-cultural backgrounds respond to this assessment? What are the consequences resulting from the intended use of the assessment results?

We question whether any engineering education research, quantitative or qualitative, should be considered positivist. Given the subjective nature of human experience, any research on human subjects needs to account for this subjectivity. As stated by Suzuki, at al. (2021), "numbers are never neutral, because they are used by humans, and therefore filtered through human biases" (p. 538). Viewing quantitative research from an interpretivist (or Critical or "post") perspective will provide richer, more complete understandings.

Funding was provided by the U.S. National Science Foundation through grants EEC-2204970, 2204971, and 2204972.

REFERENCES

AACU. 2014. “Quantitative Literacy VALUE Rubric.” https://www.aacu.org/value/rubrics/quantitative-literacy.

AERA, APA, and NCME. 2014. *Standards for Educational and Psychological Testing*. Washington, D.C: American Educational Research Association.

Baillie, Caroline, and Elliot P. Douglas. 2014. "Confusions and Conventions: Qualitative Research in Engineering Education." *Journal of Engineering Education* 103 (1):1-7. doi: https://doi.org/10.1002/jee.20031.

Cronbach, Lee J., and Paul E. Meehl. 1955. “Construct Validity in Psychological Tests.” *Psychological Bulletin* 52 (4): 281–302. https://doi.org/10.1037/h0040957.

Crotty, Michael. 1998. *The Foundations of Social Research: Meaning and Perspective in the Research Process*: Routledge.

Douglas, Elliot P., Mirka Koro-Ljungberg, and Maura Borrego. 2010. "Challenges and Promises of Overcoming Epistemological Partiality: Advancing Engineering Education Through Diverse Ways of Knowing." *European Journal of Engineering Education* 35 (3):247-257. doi: https://doi.org/10.1080/03043791003703177.

Douglas, Kerrie Anna, and Şenay Purzer. 2015. “Validity: Meaning and Relevancy in Assessment for Engineering Education Research.” *Journal of Engineering Education* 104 (2): 108–18. https://doi.org/10.1002/jee.20070.

ETS. 2021. "HEIghten Quantitative Literacy Assessment." https://www.ets.org/heighten/about/quantitative_literacy.

Fenner, Raenita, Peggy O'Neill, Kerrie Douglas, and Elliot P. Douglas. 2023. "WIP: Qualitative Content Analysis of Quantitative Literacy in First-Year Engineering Courses." ASEE Annual Conference, Baltimore, MD.

Gillman, Rick (Ed.). 2006. *Current Practices in Quantitative Literacy:* The Mathematical Association of America.

Kane, Michael T. 2015. “Explicating Validity.” *Assessment in Education: Principles, Policy & Practice* 23 (2): 198–211. https://doi.org/10.1080/0969594X.2015.1060192.

Koro-Ljungberg, Mirka, and Elliot P. Douglas. 2008. "State of Qualitative Research in Engineering Education: Meta-analysis of JEE Articles, 2005-2006." *Journal of Engineering Education* 97:163-175. doi: https://doi.org/10.1002/j.2168-9830.2008.tb00965.x.

Koro-Ljungberg, Mirka, Diane Yendol-Hoppey, Jason Jude Smith, and Sharon B. Hayes. 2009. "(E)pistemological Awareness, Instantiation of Methods, and Uninformed Methodological Ambiguity in Qualitative Research Projects." *Educational Researcher* 38 (9):687-699. doi: 10.3102/0013189x09351980.

Kosko, Karl W., and Jesse L. M. Wilkins. 2011. "Communicating Quantitative Literacy: An Examination of Open-Ended Assessment Items in TIMSS, NALS, IALS, and PISA." *Numeracy* 4 (2). doi: http://dx.doi.org/10.5038/1936-4660.4.2.3.

Mayes, Robert L., Franziska Peterson, and Rachel Bonilla. 2013. “Quantitative Reasoning Learning Progressions for Environmental Science: Developing a Framework.” *Numeracy* 6(1). doi: http://dx.doi.org/10.5038/1936-4660.6.1.4.

Messick, Samuel. 1989. “Meaning and Values in Test Validation: The Science and Ethics of Assessment.” *Educational Researcher* 18 (2): 5–11. https://doi.org/10.3102/0013189X018002005.

Messick, Samuel. 1998. “Test Validity: A Matter of Consequence.” *Social Indicators Research* 45 (1): 35–44. http://www.springerlink.com/index/X472RL52ML427811.pdf.

Mislevy, Robert J. 1994. “Evidence and Inference in Educational Assessment.” *Psychometrika* 59 (4): 439–83. https://doi.org/10.1007/BF02294388.

Mislevy, Robert J., Russell G. Almond, and Janice F. Lukas. 2003. "A Brief Introduction to Evidence-centered Design." *ETS Research Report Series* 2003 (1):i-29. doi: https://doi.org/10.1002/j.2333-8504.2003.tb01908.x.

National Academies of Sciences, Engineering, and Medicine, Division of Behavioral and Social Sciences and Education, Board on Science Education, Board on Behavioral Sciences Cognitive, and Sensory, and Committee on How People Learn II: The Science and Practice of Learning. 2018. *How People Learn II: Learners, Contexts, and Cultures*. National Academies Press.

OECD. 2012. *Literacy, Numeracy and Problem Solving Skills in Technology-Rich Environments: Framework for the OECD Survey of Adult Skills. OECD. https:doi.org/10.1787/9789264128859-en.*

Oliveri, María Elena, Jessica Nastal, and David Slomp. 2020. “Reflections on Equity-Centered Design.” *ETS Research Report Series* 2020 (1): 1–11. https://doi.org/10.1002/ets2.12307.

Pellegrino, James. 2012. “Assessment of Science Learning: Living in Interesting Times.” *Journal of Research in Science Teaching* 49 (6): 831–41. https://doi.org/10.1002/tea.21032.

Pellegrino, James. 2013. “Proficiency in Science: Assessment Challenges and Opportunities.” *Science* 340 (6130): 320–23. https://doi.org/10.1126/science.1232065.

Prince, Robert, and Zach Simpson. 2016. "Quantitative Literacy Practices in Civil Engineering Study: Designs for Teaching and Learning." 5th International Conference on Designs for Learning, Aalborg Universitetsforlag.

Riconscente, Michelle M., Robert J. Mislevy, and Seth Corrigan. 2016. "Evidence-centered Design." In *Handbook of test development*, edited by Suzanne Lane, Mark R. Raymond and Thomas M. Haladyna, 40-63. New York: Routledge.

Roohr, Katrina Crotts, Edith Aurora Graf, and Ou Lydia Liu. 2014. "Assessing Quantitative Literacy in Higher Education: An Overview of Existing Research and Assessments With Recommendations for Next-Generation Assessment." *ETS Research Report Series* 2014 (2):1-26. doi: https://doi.org/10.1002/ets2.12024.

Smith, J.J. 1989. “The Formation and Expression of Judgements in Educational Assessment.” *Journal of Further and Higher Education* 13 (3): 115–19. https://doi.org/10.1080/0309877890130310.

Songer, Nancy Butler, and Maria Araceli Ruiz-Primo. 2012. “Assessment and Science Education: Our Essential New Priority?” *Journal of Research in Science Teaching* 49 (6): 683–90. https://doi.org/10.1002/tea.21033.

Sons, Linda R. 1996*. Quantitative Reasoning for College Graduates: A Complement to the Standards*. Mathematical Association of America. https://www.maa.org/programs/faculty-and-departments/curriculum-department-guidelines-recommendations/quantitative-literacy/quantitative-reasoning-college-graduates.

Suzuki, Sara, Stacy L. Morris, and Sara K. Johnson. 2021. "Using QuantCrit to Advance an Anti-Racist Developmental Science: Applications to Mixture Modeling." *Journal of Adolescent Research* 36 (5):535-560. doi: https://doi.org/10.1177/07435584211028229

Thorndike, Robert L. 1997. *Measurement and Evaluation in Psychology and Education .* 6th ed. Upper Saddle River, NJ: Prentice Hall. http://psycnet.apa.org/psycinfo/1962-03637-000.

Vacher, H. L. 2014. “Looking at the Multiple Meanings of Numeracy, Quantitative Literacy, and Quantitative Reasoning.” Numeracy 7(2): Article 1. doi: http://dx.doi.org/10.5038/1936-4660.7.2.1.

Wankat, Phillip C., Richard M. Felder, Karl A. Smith, and Frank S. Oreovicz. 2002. “The Scholarship of Teaching and Learning in Engineering.” *Disciplinary Styles in the Scholarship of Teaching and Learning: Exploring Common Ground*, 217–37. http://fresnostate.edu/academics/csalt/documents/scholarshipofteachingandlearninginengineering-1.pdf.

Wilkins, Jesse L. M. 2010. "Modeling Quantitative Literacy." *Educational and Psychological Measurement* 70 (2):267-290. doi: https://doi.org/10.1177/0013164409344506.

Zahner, Doris, Dirk Van Damme, Roger Benjamin, and Jonathan Lehrfeld. 2021. "Measuring the Generic Skills of Higher Education Students and Graduates: Implementation of CLA+ International." In *Assessing undergraduate learning in psychology: Strategies for measuring and improving student performance.*, 219-241. Washington, DC, US: American Psychological Association.