The Deeply Engrained Behaviourist Assessment Ideologies Constraining School Mathematics

Rebecca Burtenshaw
University of the Sunshine Coast
rebecca.burtenshaw@research.usc.edu.au

Behaviourism proposes successful learning to be dependent on the performance of conditioned behaviours that are distinctly observable and objectively measurable. Over the past 100 years, various behaviourist concepts have been superseded by sociocultural and cognitive learning theories, but the entwined areas of assessment in mathematics education have not received the same focus. Outdated beliefs and unquestioned norms surrounding behaviourist assessment ideologies continue to dominate schools' decision-making around mathematics education. This paper includes the theoretical foundations for a PhD that aims to surface the tension of what is said to be valued in mathematics as "success" and what is systematically labelled as "success".

While behaviourism is theoretically viewed as obsolete, this learning theory is still inadvertently tethered to mathematics education by outdated yet deeply engrained beliefs around assessment (Shepard, 2000). The perpetuating endurance of these beliefs is a major contributor suppressing more current understandings of good mathematical learning, especially as assessment and learning should be deeply intertwined (Shepard, 2000). Musial (2021) writes how the word "assess" derives from the Latin verb *assidere*, which means to "sit beside". To "sit beside" conjures mental images of two people sitting alongside each other, determining what a learner does know or does not know yet (Musial, 2021). This is a stark contrast to assessment practices observed in some mathematics classrooms. For example, rows of individual students quietly circling answers to sets of multiple-choice questions posed to validate the attainment of normalised benchmarks (Watt, 2005). While Musial (2021) explains how assessment is commonly linked to other concepts around reporting, grades, and performance, "they are by no means the same" (p. 50). However, these concepts can all be quite synonymous with behaviourism. This paper questions if unwillingness to let go of behaviourist assessment ideologies is restraining the advancement of school mathematics education and narrowing beliefs around "success". While research into mathematics education can rarely be linked to any one theory, theoretical frameworks can act as "virtual reality systems that help practitioners connect to day-to-day realities" (Stoilescu, 2016, p.140). Therefore, this paper will explore behaviourism and how its appropriation into educational practices influenced learning, teaching and, most of all, assessment. This paper will then propose—despite time passed and research espousing sociocultural practices—that behaviourist assessment ideologies are still deeply engrained in beliefs concerning assessment and, therefore, drives behaviourist approaches in mathematics classrooms.

The Emergence of Behaviourism

Stemming from psychology, the theory of behaviourism proposes that human behaviour can be systemically conditioned to respond to specific stimuli and, if trained correctly, these behaviours are expected to be objectively observed through performance (Eisenberg, 1975; Hatfield, 2003; Stoilescu, 2016). Behaviourists suggest conditioned behaviours are adjusted through repetition, reinforcement and continuous feedback on which stimulus responses are correct and which are incorrect (Skinner, 1985; Stoilescu, 2016). Hatfield (2003) explains how the rise of behaviourism was not initially due to its own theoretical merit but instead driven by scepticism towards introspection and the knowing mind. Inspired by earlier philosophers, such as Pavlov, founding behaviourist theorists like psychologist John Watson were driven to establish that human function is measured solely by external behaviours, rejecting any influence of internal states of mind,

(2023). In B. Reid-O'Connor, E. Prieto-Rodriguez, K. Holmes, & A. Hughes (Eds.), *Weaving mathematics education research from all perspectives*. *Proceedings of the 45th annual conference of the Mathematics Education Research Group of Australasia* (pp. 123–130). Newcastle: MERGA.

emotions, and psychological conditions (Hatfield, 2003). Watson initially experimented with theories on rats, rabbits, and monkeys before applying these same experimental theories to the study of humans, which he shared in the prominent *Behaviorist Manifesto* (1913). Other behaviourist theorists at least acknowledged the presence of emotional factors, introspection, and the impact of external factors but then ultimately disregarded these in favour of externalised and measured performance (Hatfield, 2003).

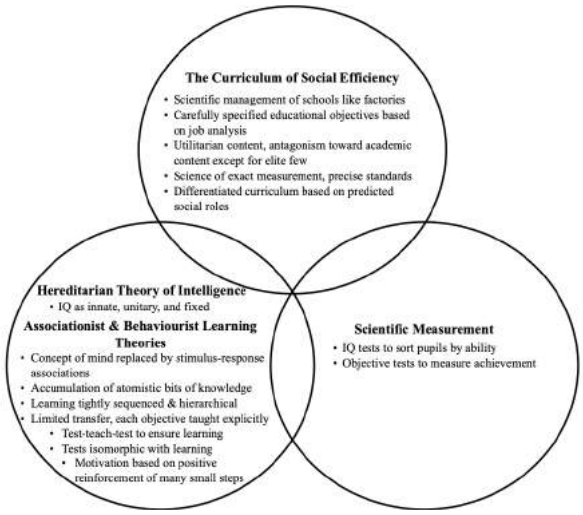

Figure 1. The interlocking paradigms that dominated the 20th Century, as adapted from Shepard (2000).

At the turn of the 20th Century, around the same time that psychology became interested in behaviour, systemic and structural developments were required to cope with population growth, movement, and revised social agendas (Hatfield, 2003; Schneider & Hutt, 2013). As seen in Figure 1, Shepard (2000) demonstrates the interlocking paradigms that dominated the 20th Century and how the overarching social efficiency of this time was closely linked to hereditarian theories of intelligence, commonly associated with behaviourism and scientific measurement. She writes how "the social efficiency movement grew out of the belief that science could be used to solve the problems of industrialisation and urbanisation" (Shepard, 2000, p. 4). Most of the social efficiency movement involved some form of depersonalisation and standardisation (Shepard, 2000), which educational institutions began to rely upon to maximise school efficiency (Schneider & Hutt, 2013). For example, Schneider and Hutt (2013) write how "administrators refashioned themselves as professional managers whose job was to manage burgeoning systems in the most efficient way possible" (p. 7). Reporting systems that relied upon uniform standards and easily communicated grades were modified from other industries, such as cattle and timber (Schneider & Hutt, 2013). Another example of an incipient measurement scheme was the Intelligence Quotient (IQ), which attempts to measure someone's intellectual potential or lack of (Gipps & Stobart, 2009). Edward Thorndike, credited as the "father of scientific measurement" (Shepard, 2000) and a founder of educational psychology, was a strong supporter of the emerging IQ assessment. Alongside his studies in human conditioning, Thorndike (1910) believed an individual's original nature—intellectually, characteristically, and morally—is inherently bound by ancestry, and individuals should be educated according to their innate genetic capabilities. Though, he also believed a "man's original nature" could be improved and that this conditioning would contribute to the "success in controlling human nature and changing it to the advantage of the common weal" (Thorndike, 1910, p. 8). Hilgard (1996) writes how Thorndike and John Dewey were colleagues for many years but held different views about education. While Dewey was a progressive reformer, Thorndike "valued data above all else" (Hilgard, 1996, p. 423) and believed school improvement was achieved through quality control rather than innovation. The education system favoured and embraced Thorndike's theories as it accommodated other beliefs at this time. Instigated by various societal factors,

behaviourism learning theory emerged and was applied to various school subject areas, such as mathematics education.

Behaviourism in Mathematics Education

Behaviourism learning theory commonly characterises learning and instruction as the transmission of carefully sequenced knowledge passed down from one to another and typically accomplished through the mimicking and repetition of universal, static procedures and instructed strategies (Askew et al., 1997; Richardson, 1996; Stoilescu, 2016). Stoilescu (2016) explains how behaviourism is commonly intertwined with "the belief that mathematics can be transmitted by inoculating the right knowledge and discourse at the right time" via "already established theories, knowledge paths and experiences" (p.141). Behaviourism, the oldest and most well-known theoretical orientation, aligns closely with a transmission model (Richardson, 2001; Stoilescu, 2016). Traditionally, behaviourism learning theory centres teachers as the arbiter of knowledge (Richardson, 2001). However, more modern representations of this learning theory include technology programs platforms playing a role in instruction and the transmission of knowledge (Aydin, 2005). With each mathematical concept sequenced within a wider hierarchy, students are situated as the passive recipients of knowledge previously discovered and owned by masters centuries ago (Bada & Olusegun, 2015; Handal, 2009). This style of learning is often completed through explicit procedural teaching, drill-like practice, and repetitive rote learning (Askew et al., 1997; Handal, 2009). Such an approach also directly aligns with behaviourist beliefs around conditioning behaviour. In fact, within both a transmission model and behaviourism learning theory, each lesson will often have pre-determined goals with pre-determined outcomes, requiring achievement in pre-determined ways (Shepard, 2000). Therefore, "success"—for both the teacher and the student—is judged (or "assessed") on how well a student can repeat and perform the desired behaviour (Jacob et al., 2017).

Behaviourism learning theory views assessment of learning as students' displaying quantifiable changes in behaviours that are distinctly observable and objectively measurable (Eisenberg, 1975; Stoilescu, 2016). This objectivity is repetitively noted as a superior trait that should drive the adoption of behaviourist ideologies (Eisenberg, 1975; Stoilescu, 2016). An example in mathematics education is short answer exams that examine the accuracy of followed procedures to arrive at an answer. The number of correct responses is scored against an overall total and is often then related to a grade. This objective 'assessment' is frequently accompanied by themes of scientific measurement, such as norms, ranking and grading, and terms like 'evidence', 'standards' and 'outcome-based education' (Eisenberg, 1975; Maxwell, 2009; Shepard, 2000; Stoilescu, 2016). It is proposed these "precise standards of measurement" are required to validate that learning has occurred and that each skill is "mastered at the desired level" (Shepard, 2000, p. 4). As Thorndike (1910) claims, if education is about change, then there should be units of measurement to observe such change. While this may be true in some applications, like most things in education, balance is needed when considering any initiative or educational practice.

Some behaviourist and hereditarian theories of intelligence have since been disputed in education, as well as psychology. For example, statements made by Thorndike (1910) have been empirically challenged, such as the accuracy of IQ measurement (Gipps & Stobart, 2009), the varying intellectual capacity of different "races", and the superior mental function of males compared to females. Furthermore, the revolution of cognitive research, social learning theory and Postmodernism dispositions confronted behaviourist theories (Richardson, 2001). Research that supports the ongoing impact of introspection (Desautel, 2009; Munns & Woodward, 2006), the influence of a growth mindset (Dweck, 2014) and the ongoing developments founded through cognitive science. A wide range of research also claims the harms of behaviourist ideologies. For example, Eisenberg's (1975) paper, *Behaviourism: The Bane of School Mathematics*, shared

numerous examples of the damaging impact behaviourism can have on mathematical understanding. Piaget and Vygotsky's theories around cognitive development have inspired many more sociocultural learning theories (Goos, 2004; Stoilescu, 2016). In teacher education, behaviourism is commonly noted as a theory of the past in favour of more current learning orientation, such as constructivism (Richardson, 1996). Constructivism orients learning as an active process of meaning-making developed by and adapted from interactions with others, experiences (prior and new) and their subjective states of mind (Bada & Olusegun, 2015; Handal, 2009; Stoilescu, 2016). Teachers as facilitators provide adaptive and experiential learning opportunities as students create their own 'personal mental model' of understanding (Bada & Olusegun, 2015) and "look for similarities and differences against their own cognitive schemata" (Handal, 2009, p. 5).

There are many examples of literature (Bada & Olusegun, 2015; James, 2006; Richardson, 1996, 2001) that discusses both behaviourism and constructivism, typically beginning with a denunciation of behaviourism followed by an exploration of constructivism. This article does not intend to compare or pit behaviourism against constructivism. In fact, Aylward and Cronjé (2022) contend that behaviourism and constructivism should not be viewed as linear opposites as traditionally posed but complementary. However, this behaviourism-constructivism flow of discussion could unintentionally lead the reader to interpret that one has replaced the other. A careful distinction needs to be made about how learning orientations (like behaviourism and constructivism) continue to exist side by side. While one learning orientation may have a greater preference or be more influential in particular subject areas, from its inception, learning orientations exist in parallel (Stoilescu, 2016). It is important to acknowledge that school learning and teaching practices may have evolved to embed cognitive learning ideologies, constructivist frameworks and other sociocultural theories espoused by mathematics education research (James, 2006; Shepard, 2000). However, there is still a strong presence of behaviourist ideologies in mathematics assessment practices. Shepard (2000) writes how "dominant theories of the past continue to operate as the default framework affecting and driving current practices and perspectives" (p.4). These out-of-date ideas can still shape the belief systems of teachers, students, parents, and policymakers (Shepard, 2000), though no more prominently than within assessment.

There is a deep obligation to the behaviourist's notion of objectivity within mathematics education assessment, even if this attachment to psychometric measurement comes at the detriment of student learning and other important assessment design principles (Shepard, 2000; Watt, 2005). Watt (2005) investigated Sydney mathematics teachers' attitudes towards alternative assessment methods in mathematics and found an astounding overreliance on written tests, despite participants also indicating how "traditional mathematics tests cannot be used to infer more general mathematical ability" (p. 23). Participants across various years of teaching experience indicated that alternative assessments—even quite a conservative shortlist, such as student self-assessment, oral tasks and practical tasks—were perceived as too subjective and, in a lesser sense, unsuitable for mathematics (Watt, 2005). The perceived "threat to the objectivity" (Watt, 2005, p. 39) may override the rationality to collate a well-rounded, holistic picture of students' mathematical understanding, which can not fairly or ethically achieved from one assessment modality (Clarke, 1997). Objectivity is also highly valued when assessment is exclusively connected to grading (Watt, 2005). However, Clarke (1997) writes, "Assessment is a process. Grading can be one product of that process. The two should not be confused with each other" (p. 21). Viewing grading and assessment as indistinguishable misrepresents and underestimates the value of assessment and its broader application in classrooms (Clarke, 1997; Munns & Woodward, 2006). After all, "no one has ever gotten taller just by being measured" (Clarke, 1997, p. 4).

That is not to say objectivity is unimportant; however, other rigorous assessment design principles, such as validity, clarity, fidelity, and fairness, are disproportionately considered in the pursuit of objectivity (Musial, 2021; Zane, 2009). Watt (2005) observes, "To date, reliance on the

traditional mathematics test has been justified on the grounds of maximising reliability and ensuring comparability, but this has often been at the expense of validity" (p. 24). Validity refers to the extent evidence and design measure what was intended to be measured (Rawlins et al., 2005). Gipps and Stobart (2009) also believe fairness and equity should be embedded within validity arguments by examining sociocultural dispositions, such as the type of knowledge valued, assessment preparation, how assessment is used and the ramifications of such use. They believe the discussion around validity should move from "a *fixed property* of an assessment" to a "*process* that investigates an assessment in terms of both the construct being assessed… and, crucially, the inferences and actions based on the results" (p. 109-110). Assessment that optimises students' expression of their learning requires purposeful consideration of a variety of assessment design principles, including transparency about the measures of success, how this judgment is reached and follow-up actions that continue to support student learning (Clarke, 1997; Rawlins et al., 2005). Without this, the single-minded obligation to objectivity may lead to invalid judgments, restrictive opportunities, and the risk of misinterpreting student learning (Clarke, 1997; Rawlins et al., 2005).

There are many other examples of behaviourist assessment ideologies that are widely implemented and deeply engrained in mathematics classrooms, such as:

- The unceasing teach-test-teach-test culture which requires students to demonstrate their atomised and, possibly, shallow understanding of a specific concept after short-term memorisation (Eisenberg; 1975, Shepard, 2000), potentially resulting in the same concepts being retaught year after year.
- The ongoing consequences of NAPLAN and the ill-conceived notion that high-stakes standardised testing is more reliable than the judgement of education professionals (Klenowski & Wyatt-Smith, 2012). This includes schooling systems using NAPLAN scores to reward or reprimand schools (Thompson & Cook, 2014) and, as Watt (2005) describes, the "backwash" of such requirements and pressure.
- Common practices of collecting "data for data's sake", with often limited opportunity to enact upon or give feedback to students (Black & Wiliam, 2004; Liljedahl, 2021). This also includes poor implementation of this data and the use of data to label or rank students by numerical or abstract grading bands (Clarke, 1997; Watt, 2005).
- The streaming of students and classes (Tieso, 2003).
- Inflexible learning objectives attached to transmission models of mathematics teaching, including pushing through the curriculum and limited differentiation (Shepard, 2000).
- The frequent use of the term "standards", despite the real possibility that completely incomparable stances exist (Maxwell, 2009). Maxwell (2009) explains how standards may differ depending on the type of standard, the focus (facet, unit, scope), the construct and the purpose. For example, the different constructs between determining learning and performance; or the distinct implementations of content standards, performance standards and development standards.
- Schooling systems commending theories regarding measurement models, such as Hattie's (1992) effect sizes. This can reinforce the importance of measurement models in schools and certain pedagogical decision-making (Serow et al., 2016).
- Marginalised representations of what "knowledge" is and which areas of knowledge are deemed as more important. (Gipps & Stobart, 2009; Seth, 2007).

In some of these examples, the limitations and constraints may be acknowledged by schools and educators but ultimately disregarded due to normalisation (Klenowski & Wyatt-Smith, 2012; Serow et al., 2016). However, in other cases, the saturation of practices may be so systemic that the influences of behaviourist assessment ideologies are overtly unrecognisable.

While research has questioned what we think we know about teaching practices and student learning, it is important to ask similar questions about assessment ideologies. If we do not do this, the developments of mathematics education will be in vain as the prevailing dominance of behaviourist assessment ideologies will continue to constrain the holistic forward movement of mathematics education. For example, if decision-making around mathematics education practices is made to match the behaviourist assessment norms and requirements, then there may be hesitation in adopting more current researched practices. This could include teachers' unease in exploring the role of metacognition in mathematical thinking (Desautel, 2009) or providing an environment for students to engage together in creative problem-solving play (Liljedahl, 2021) for fear of the time taken away from outcome-based objectives. Furthermore, if mathematics learning and teaching decisions are only made to match out-of-date assessment norms, then similarly out-of-date practices will always be perceived as more robust and fitting (Shepard, 2000). If assessment only requires students to repeat the thinking of others, then teaching will primarily focus on students learning *what* to think, as opposed to *how* to think. "Assessment should model the mathematical activity we value", writes Clarke (1997, p. 8).

Broader Implications

Questioning the depth of behaviourist assessment ideologies has wider implications beyond assessment, including belief systems for students and teachers, approaches to mathematics education and the motivations for engaging with other mathematics education research (Beswick, 2012; Ekmekci et al., 2015; Stoilescu, 2016). Shepard (2000) explains that if there is to be "any attempt to change the form and purpose of classroom assessment", including making it a more interconnected part of learning, then we "must acknowledge the power of these enduring and hidden beliefs" (p. 6). For example, Ernest (1989) describes three categories of beliefs influencing mathematics: Instrumentalist, Platonist and Problem Solving. There is a tangible congruence between behaviourism and traditional Instrumentalist philosophies of mathematics that could contribute to the continuation of conservative beliefs regarding the nature of mathematics (Ernest, 1989; Handal, 2009; Hatfield, 2003). Mendick (2005) alleges that this view of mathematical knowledge "as absolute and unquestionable" creates the status of "the ultimate intelligence test". This can contribute to the sustained elitism, irrelevance and disaffection towards mathematics, which Nardi and Steward (2003) describe. Richardson (2001) explains how the late 1960s and early 1970s saw research on teaching emerge as its own field of study with an emphasis on process-product research. Process-product research focused on finding correlations between teacher behaviours and improved student achievement (such as standardised test scores) to gauge the difference between more or less 'effective teaching' (Richardson, 2001). Firmly established studies, such as direct instruction in mathematics, were conducted under the belief that teachers were the transmitters of knowledge and were influenced by behaviourist ideologies (Richardson, 2001). Though the notions of an "effective teacher" shifted from teacher behaviours to "teacher cognition, beliefs and knowledge" (Richardson, 2001, p. 282), these initial ideas of 'effective teaching' constructed a strong image of what it means 'to teach'. Jacob et al. (2017) also demonstrate how students' performance on tests and assessment are still deeply interconnected with teachers' beliefs around the success of their own teaching and what it means to be "a good teacher". Students' perceived and realistic successes and failures emotionally drive teachers' own emotions, interactions, and pedagogical decisions (Jacob et al., 2017). Therefore, beliefs intertwined with behaviourist assessment ideologies influence beyond assessment.

Back in 1975, Eisenberg spoke about the inevitable damage and obscured "gravity of espousing an extreme behaviouristic format" and how—even then—"acting under the guise of accountability, [has] forced classroom teachers and curriculum developers via legislation and funding policies to adopt a behaviouristic framework" (p. 163). Further investigation is required to understand the extent, scope and repercussions of how engrained behaviourist assessment ideologies are, including

how assessment could challenge or reinforce the broader beliefs of mathematics. By valuing process equally to or more than product (Handal, 2009; Stoilescu, 2016; Watt, 2005), the "assessment of success" in mathematics could shift perceptions, break down limitations and unlock the subject of mathematics—be it in school, careers, or daily life. However, first, the cyclical perpetuation of behaviourism ideologies stemming from and sustained through assessment will need to be surfaced and questioned. These are the hopes and objectives which my PhD aims to address. There is an ambiguity between what is said to be valued in mathematics education and labelled as "success", and what is systemically demonstrated as valued or judged as "success". Surfacing this tension will reveal beliefs around "success" in mathematics education and begin to broaden success metrics beyond traditional behaviourist structures ingrained in Australian education.

References

Askew, M., Rhodes, R., Brown, M., Wiliam, D., & Johnson, D. (1997). *Effective teachers of numeracy*. King's College London. http://musicmathsmagic.com/page2/files/EffectiveTeachersofNumeracy.pdf

Aydin, E. (2005). The use of computers in mathematics education: A paradigm shift from "computer assisted instruction" towards "student programming". *Turkish Online Journal of Educational Technology*, *4*(2), 4.

Aylward, R. C., & Cronjé, J. C. (2022). Paradigms extended: How to integrate behaviorism, constructivism, knowledge domain, and learner mastery in instructional design. *Educational Technology Research and Development*, *70*(2), 503–529. https://doi.org/10.1007/s11423-022-10089-w

Bada, S. O., & Olusegun, S. (2015). Constructivism learning theory: A paradigm for teaching and learning. *Journal of Research & Method in Education*, *5*(6), 66–70.

Beswick, K. (2012). Teachers' beliefs about school mathematics and mathematicians' mathematics and their relationship to practice. *Educational Studies in Mathematics*, *79*(1), 127–147. https://doi.org/10.1007/s10649-011-9333-2

Black, P., & Wiliam, D. (2004). The formative purpose: Assessment must first promote learning. *Yearbook of the National Society for the Study of Education*, *103*(2), 20–50. https://doi.org/10.1177/016146810410601402

Clarke, D. (1997). *Constructive assessment in mathematics: Practical steps for classroom teachers*. Key Curriculum Press.

Desautel, D. (2009). Becoming a thinking thinker: Metacognition, self-reflection, and classroom practice. *Teachers College Record*, *111*(8), 1997–2020. https://doi.org/10.1177/016146810911100803

Dweck, C. (2014). Teachers' mindsets: "Every student has something to teach me." Feeling overwhelmed? Where did your natural teaching talent go? Try pairing a growth mindset with reasonable goals, patience, and reflection instead. It's time to get gritty and be a better teacher. *Educational Horizons*, *93*(2), 10–15. https://doi.org/10.1177/0013175X14561420

Ekmekci, A., Corkin, D., & Papakonstantinou, A. (2015). The relationship between teacher related factors and mathematics teachers' educational beliefs about mathematics. In S. M. Che, & K. A. Adolphson (Eds.), *Proceedings of the 42nd annual meeting of the Research Council on Mathematics Learning* (pp. 140–148).

Eisenberg, T. A. (1975). Behaviourism: The bane of school mathematics. *International Journal of Mathematical Education in Science and Technology*, *6*(2), 163–171. https://doi.org/10.1080/0020739750060204

Ernest, P. (1989). The impact of beliefs on the teaching of mathematics. In P. Ernest (Ed.), *Mathematics teaching: The state of the art* (pp. 249–253). Falmer.

Gipps, C., & Stobart, G. (2009). Fairness in assessment. In C. Wyatt-Smith, & J. J. Cummings (Eds.), *Educational assessment in the 21st century: Connecting theory and practice* (pp. 105–118.) Springer Science + Business Media. https://doi.org/10.1007/978-1-4040-9964-9_6

Handal, B. (2009). Philosophies and pedagogies of mathematics. *Ilköğretim Online*, *8*(1), 1–6.

Hatfield, G. (2003). Behaviourism and psychology. In T., Baldwin. *The Cambridge history of philosophy* 1870–1945. Cambridge University Press. https://doi.org/10.1017/CHOL9780521591041

Hattie, J. (1992). Measuring the effects of schooling. *Australian Journal of Education, 36*(1), 5–13.

Jacob, B., Frenzel, A. C., & Stephens, E. J. (2017). Good teaching feels good—but what is "good teaching"? Exploring teachers' definitions of teaching success in mathematics. *ZDM Mathematics Education*, *49*(3), 461–473. https://doi.org/10.1007/s11858-017-0848-6

James, M. (2006). Assessment, teaching and theories of learning. In J. Gardner (Ed.), *Assessment and learning*, (pp. 47–60). Athenaeum Press.

Klenowski, V., & Wyatt-Smith, C. (2012). The impact of high stakes testing: The Australian story. *Assessment in Education: Principles, Policy & Practice*, *19*(1), 65–79. https://doi.org/10.1080/0969594X.2011.592972

Liljedahl, P. (2021). Build*ing thinking classrooms in mathematics: 14 teaching practices for enhancing learning: Grades K–12*. Corwin.

Maxwell, G. S. (2009). Defining standards for the 21st century. In C. Wyatt-Smith, & J. Cumming (Eds.), *Educational assessment in the 21st century* (pp. 263–286). Springer.

Mendick, H. (2005). Mathematical stories: Why do more boys than girls choose to study mathematics at AS-level in England? *British Journal of Sociology of Education*, *26*(2), 235–251.

Munns, G., & Woodward, H. (2006). Student engagement and student self-assessment: The real framework. *Assessment in Education: Principles, Policy & Practice*, *13*(2), 193–213.

Musial, D. L. (2021). Meaningful assessment. In L. Wellner, & K. Pierce-Friedman (Eds.), *Supporting early career teachers with research-based practices* (pp. 50–70). IGI Global.

Nardi, E., & Steward, S. (2003). Is mathematics TIRED? A profile of quiet disaffection in the secondary mathematics classroom. *British Educational Research Journal*, *29*(3), 345–367. https://doi.org/10.1080/01411920301852

Rawlins, P., Brandon, J., Chapman, J., Leach, L., Neutze, G., Scott, A., & Zepke, N. (2005). Standards-based assessment in the senior secondary school: A review of the literature. *New Zealand Journal of Teachers' Work*, *2*(2), 107–115.

Richardson, V. (1996). From behaviorism to constructivism. *Teacher Education and Special Education*, *19(*3), 263–271.

Richardson, V. (2001). Constructivist mathematics instruction and current trends in research on teaching. In T. Wood, B. Scorr Nelson, & J. Warfield (Eds.), *Beyond classical pedagogy: Teaching elementary school mathematics* (pp. 275–293). Routledge.

Schneider, J., & Hutt, E. (2014). Making the grade: A history of the a–f marking scheme. *Journal of Curriculum Studies*, *46*(2), 201–224. https://doi.org/10.1080/00220272.2013.790480

Seth, S. (2007). Changing the subject: Western knowledge and the question of difference. *Comparative Studies in Society and History*, *49*(3), 666–688. https://doi.org/10.1017/S0010417507000667

Serow, P., Callingham, R., & Tout, D. (2016). Assessment of mathematics learning: What are we doing? In K. Makar, S. Dole, J. Visnovska, M. Goos, A. Bennison, & K. Fry (Eds.), *Research in mathematics education in Australasia 2012–2015* (pp. 235–254). Springer.

Shepard, L. A. (2000). The role of assessment in a learning culture. *Educational Researcher*, *29*(7), 4–14. https://doi.org/10.3102/0013189X029007004

Skinner, B. F. (1985). Cognitive science and behaviourism. *British Journal of Psychology*, *76*(3), 291–301. https://doi.org/10.1111/j.2044-8295.1985.tb01953.x

Stoilescu, D. (2016). Aspects of theories, frameworks and paradigms in mathematics education research. *European Journal of Science and Mathematics Education*, *4(*2), 140–154.

Tieso, C. L. (2003). Ability grouping is not just tracking anymore. *Roeper Review*, *26(*1), 29–36. https://doi.org/10.1080/02783190309554236

Thompson, G., & Cook, I. (2014). Manipulating the data: Teaching and NAPLAN in the control society. *Discourse: Studies in the Cultural Politics of Education, 35*(1), 129–142.

Thorndike, E. L. (1910). The contribution of psychology to education. *Journal of Educational Psychology*, *1*(1), 5–12.

Watson, J. B. (1913). Psychology as the behaviorist views it. *Psychological Review*, *20*(2), 158–177. https://doi.org/10.1037/0033-295X.101.2.248

Watt, H. M. (2005). Attitudes to the use of alternative assessment methods in mathematics: A study with secondary mathematics teachers in Sydney, Australia. *Educational Studies in Mathematics*, *58*(1), 21–44.

Zane, T. W. (2009). Performance assessment design principles gleaned from constructivist learning theory (Part 2). *TechTrends*, *53*(3), 86. https://doi.org/10.1007/s11528-009-0287-5