

Improvement and accountability functions of assessment: Impact on teachers thinking and action

Gavin T L Brown, PhD

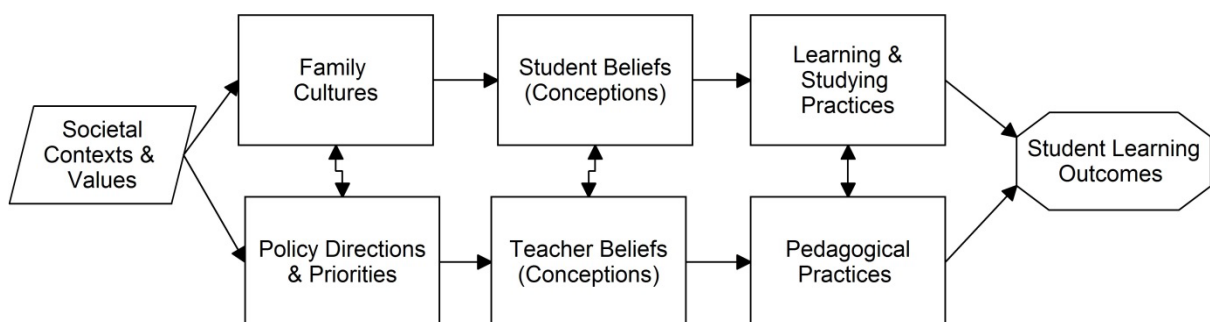
*The University of Auckland*

Assessment is everywhere in schooling. In some countries it begins before schooling starts with screening tests and interviews for selection into elite or competitive nursery schools or kindergartens. Nonetheless, for most children and adolescents, if assessment is not a regular occurrence, it looms in the background, perhaps haunting or driving their learning. Likewise, assessment plays an important role in curriculum and teaching; assessment is how teachers monitor students' progress through the curriculum and how curricula can be evaluated for effectiveness. Unsurprisingly, assessment is also used by governments, the media, and parents to determine the quality of schools and teachers; even though there is strong evidence that such uses have generally negative consequences on teaching, curriculum, students, and learning (Hamilton et al., 2007); although some positive consequences (e.g., teaching being focused more directly on the curriculum underlying the test) have been documented (Cizek, 2001).

Educational policy shapes the context in which teachers perform their multi-faceted work (e.g., planning, teaching, and evaluating). Policy expresses the societal and cultural norms valued by members of that jurisdiction. Thus, the introduction of policy reform around assessment (e.g., No Child Left Behind or National Standards vs. Assessment for Learning) may express values not necessarily held by teachers employed to implement such policies. When policy needs to be implemented by teachers, it cannot be assumed that teachers will enact policy as intended. Indeed, there is strong evidence that how teachers conceive of a phenomenon acts to filter (i.e., control what they pay attention to), frame (i.e., control how they understand what they pay attention to), and guide (i.e., influence their behaviour) their responses to that phenomenon (Fives & Buehl, 2012). Hence, attention needs to be paid to the conceptions or beliefs teachers have surrounding current assessment policies and priorities so as to better appreciate how they are likely to understand, respond to, and implement assessment reforms. This is especially the case when school systems place increasing responsibility for assessment in the hands of schools or teachers, rather than solely in the control of external examination boards. If teachers conceive that assessment is primarily about evaluating student performance or if they conceive that learning from assessment (i.e., improvement) is largely what the student is supposed to do, then it is unlikely they will use assessment information to reconsider their own teaching practices, even if students in their class do poorly on one or more topics covered by the assessment. Offering teachers more training in test statistics, test item writing skills, or curriculum content knowledge, without addressing their conceptions of what assessment is and what it is for might simply create better skilled teachers who still think assessment is about the student rather than their own practice (Brown, 2008).

The term conceptions is frequently used in examining teacher beliefs about assessment because it captures not only how assessment is perceived but also the cognitive and affective thoughts, opinions, and attitudes that teachers form about a phenomenon such as assessment. Teacher conceptions are arrived at through perceptions and embodied experiences leading to the formation of beliefs. Since societies offer different experiences, it is only to be expected that differences in culture or society lead not only to differing policies, but also to distinct conceptions of practices or processes. For example, transmission teaching coupled with high respect for teachers and formal examination scores characterizes East Asian societies (Hofstede, 2007) as a legitimate means of motivating students, rewarding diligence, and overcoming negative social effects such as promotion through collusion, corruption, or nepotism. In contrast, the Anglo-Commonwealth schooling world is characterized by strong attention to the needs and values of individual children in which the teacher facilitates each child's talents or personal priorities (Stobart, 2006). In such societies, assessments that prevent children from gaining access to further learning opportunities is frowned upon, at least by schooling professionals. Thus, it is expected that teacher beliefs within each jurisdiction or ecology will be broadly shared and coherent with the priorities of both policy and society. Under these conditions, teacher beliefs will lead to decision making that could be understood as ecologically rational within that environment (Rieskamp & Reimer, 2007). Beliefs that are ecologically rational lead to actions and priorities that make sense within the overall set of priorities of a society. This implies that in a contrasting context, such beliefs might be seen as irrational or unacceptable.

Hence, the conceptual model underlying research into teacher beliefs about assessment (Figure 1) has twin, interacting tracks leading to student outcomes; the conceptions of both teachers and students in a jurisdiction are influenced by various policy directions and family priorities and these beliefs, in turn, guide their separate teaching and learning practices. These two pathways are shaped by and respond to societal and cultural contexts, meaning that there will be different beliefs and practices in differing social, ethnic, and cultural groups. The model has three important characteristics. First, teacher beliefs moderate or mediate what happens between policy directions and student learning outcomes, rather than relying on processes external to the implementation environment. Second, policy directions are seen as a function of priorities within society and culture, suggesting that variation in conceptions and practices within societal contexts will be less than those between contexts. Third, the beliefs students themselves have about schooling, learning, teaching, knowledge, curriculum, and assessment play a strong contributing role in shaping their outcomes.



*Figure 1.* Conceptual framework of relations leading to outcomes

It is worth considering, as Pajares (1992) suggested, that teacher beliefs or conceptions arise out of their student experiences within a system. This interaction is shown in the double-headed arrows between teacher and student in the diagram. This means that it is likely that the conceptions of assessment students develop through successfully navigating assessment in schooling (i.e., passing tests or exams, doing well on performance or portfolio assessments, gaining qualifications, etc.) stimulate beliefs about the nature and purpose of assessment that student teachers bring with them upon entry to teacher education. These beliefs, of course, may not be appropriate for the role of teacher, depending on policy and practice environments.

### **Assessment Purposes**

Instead of focusing on assessment formats and types (e.g., essays, multiple-choice tests, etc.), it seems more useful to focus on the purposes teachers have for whatever assessment techniques they employ. Assessment has many uses—selection, promotion, retention, deciding awards, grouping students, certification, reporting, tracking progress, and so on (Newton, 2007). Many of these uses are largely administrative rather than educational. While administrative demands are legitimate, in and of themselves, they are often only indirectly related to the primary concerns of educators. Educational uses of assessment focus much more on the possibility that assessment can inform improved student learning and better teaching.

Popham (2000, p. 1) eloquently and forcefully stated:

if educational measurement doesn't lead to better education for students, then we shouldn't be doing it ... the only reason educators ought to assess students is in order to make more defensible educational decisions regarding those students. That's really why educators should be messing around with measurement-*to improve student learning*. [italics in original]

While Popham uses the term measurement, we can infer that he means assessment—a process of collecting information about student learning that leads to educational decision making. The key educational decision that every teacher needs to engage in is deciding *who needs to be taught what next*. While educational improvement is an essential goal of assessment, this ambition is often overwhelmed by not just administrative, but also accountability uses, normally imposed by political authorities.

Accountability has to do with a simple idea: everyone has to give an account of what they have done in their work. In education, teachers are accountable for their students' learning and they are required to account for their effectiveness to their managers, leaders, and supervisors and, naturally, the parents of their students. A simple way, though not without problems, to evaluate teachers is to test their students—if teachers have done a good job, their students will do well, or at least better than they did previously, on any assessment of the curriculum teachers were supposed to teach. However, there are many reasons students might

not do well on an assessment, independent of how well the teacher has done his or her job. For example, the teacher may be working with students from poor homes, whose parents have little education, or use a language different to that of the school. Alternatively, even an average teacher may appear very good, if the home resources of his or her students are high and those students are extensively helped to learn in out of school tutoring for the assessment.

What makes accountability assessment important is that there are often negative consequences for schools and teachers who receive low scores. In examination driven societies, student assessment results are published in the media and used to determine the worth or quality of schools and teachers. In extreme cases, teachers have been fired, students have been forced to repeat grades, and schools have been disbanded. Being associated with high or low achieving schools brings reflected glory or shame to teachers. However, accountability that depends on tests or assessments is a two-edged sword. On the one hand, higher scores are meant to indicate that students have achieved what society expected from schools, and, on the other hand, high scores can be inflated without ensuring that the expected learning has taken place. Teaching to and cramming for the test, let alone cheating, can result in higher scores. Evidence exists that, when teachers believe that the negative consequences associated with low scores are unfair or unethical, test score inflation practices will take place (Ravitch, 2013). These consequences are easily understood when we realise that accountability has fairly consistent and powerful effects on human—we tend to comply with the views of those to whom we must give an account of our work (Lerner & Tetlock, 1999). This means that teachers will produce results that meet what they know their superiors expect—and if the consequences of failure are extreme for the individual concerned (e.g., loss of reputation)—then we should expect teachers to deliver high test scores, through whatever means at their disposal. It may also mean that if schools are truly community-based, then teachers may disregard standardised tests completely, if that is what is expected by a parent and leadership group to whom they are responsible. Further, it may mean that teachers would ignore high scores on tests if the students in their class, to whom they are also accountable, press for a broad and diverse curriculum experience rather than a narrow teaching-to-the-test version.

### **Teacher beliefs about assessment**

Brown (2008) has framed teacher beliefs about assessment as a multi-dimensional space in which four competing purposes interact with each other; these are (1) assessment informs improved teaching and learning, (2) assessment holds schools and teachers accountable, (3) assessment holds students accountable, and (4) assessment is irrelevant or ignored. Extending this framework, assessment purposes can be positioned on a continuum from extremely pedagogical to extremely accounting (Barnes, Fives, & Dacey, 2015). Harris and Brown (2009) positioned teacher beliefs about assessment in a 2\*2 frame according to whether they were predominantly (1) about students or schools and (2) positive or negative evaluations. Generally, their analysis of 26 New Zealand teachers concluded that teachers were most in favour of assessments that improved pedagogical interactions between teachers and students and among students themselves.

Research with the Teachers' Conceptions of Assessment self-report inventory has provided insights into teacher thinking by examining both mean scores and correlations among these four major factors. International comparative studies have generally concluded that teachers give the highest level of agreement to the use of assessment for improvement or pedagogical applications. In addition to New Zealand and Queensland, the improvement purpose had highest value in jurisdictions (see summary in Brown, 2012). Consistent with the idea that teacher beliefs about assessment develop initially from their experiences and perspectives as students (i.e., assessment evaluates me the student), research studies have found that prospective teachers do not have improvement as their dominant. However, deliberate training of prospective teachers to take on a formative or improvement-oriented approach to assessment has found that future teachers do become more oriented towards pedagogical approaches to assessment (Smith, Hill, Cowie, & Gilmore, 2014).

Societies can be grouped as to whether the assessment policy focuses on high-stakes consequences (e.g., ranking of schools, entry awards for students, sanctions for poor performance, etc.) or focuses more on low-stakes consequences (e.g., diagnostic analysis of student needs, provision of support to students or teachers, etc.). An important trend is that similarity in teacher beliefs is found across jurisdictions with similar policies and cultures. For example, Hamilton, et al. (2007) reported that teachers in California, Georgia, and Pennsylvania had very similar responses, experiences, and attitudes towards standards-based accountability assessments; attributable to similarities between the systems. Likewise, in New Zealand and Queensland, which both have low-stakes, child-centred, formative assessment systems, primary school teachers had statistically equivalent responses to the Teacher Conceptions of Assessment inventory (Brown, Lake, & Matters, 2011). Similarly, teachers in Hong Kong and China, both of which are high-stakes, public examination societies, had statistically equivalent responses to a Chinese-Teacher Conceptions of assessment inventory (Brown, Hui, Yu, & Kennedy, 2011). The similarity of Hong Kong and China and their consistent differences with New Zealand and Queensland has been attributed to the Confucian-Heritage cultural features of the assessment system in which examinations are a force for both improved learning and improved personal character, while the latter two societies have a more open view in which assessment does not function as a barrier to further opportunity.

Where problems arise for teachers is when supposedly low-stakes policies (e.g., assessment for learning) are implemented alongside high-stakes examinations or school evaluation systems. In these situations, it is highly likely that formative, diagnostic approaches to assessment will be treated as a 'soft' policy, in contrast to formal examinations or school evaluation systems that function as 'hard' policy (Kennedy, Chan, & Fok, 2011). For example, although Hong Kong has an assessment for learning policy, the society is characterised by high-consequence, examinations that determine life chances for children and, by inference, determine the quality of schools and teaching. Thus, teachers strongly associate improvement with holding students accountable (Brown, Kennedy, Fok, Chan, & Yu, 2009; Brown, Hui, Yu, & Kennedy, 2011), unlike teachers in New Zealand and Queensland (Brown, 2012; Brown, Lake, & Matters, 2011). Interestingly, New Zealand and

Queensland (Brown, 2012) secondary school teachers both agreed more with the purpose that assessment holds students accountable than their primary counterparts; this consistent difference was attributed to the role secondary teachers play in assessing students for qualifications and the greater responsibility adolescents are expected to take for learning.

Hence, teacher conceptions of assessment are, in part, a product of how teachers individually experience assessment in their student careers (usually successfully through the application of individual effort and home and school resources) and how society constructs the role and function of assessment in schooling. At the same time, a consequence of teacher conceptions of assessment is a replication in their practices of the societal norms that contributed to their own success as assessed learners. In other words, *what worked for me as a student* (usually assessment evaluates me—Brown, 2008) *is how I will teach*. Since successful students do accept the legitimacy of having their learning evaluated, teachers who emphasise this conception potentially close off the questions that effective teachers should ask: *if my students do not succeed on an assessment, is it because I didn't teach them as well as they needed?* Reflective and effective teachers are open to the possibility that the assessment data are correct and that questions ought to be posed instead concerning their teaching. Good education systems support the teacher in discovering the bad news about their own teaching, rather than laying responsibility solely upon society or students.

Nevertheless, policy change around the use of assessment (e.g., increased accountability or increased formative expectations) cannot be successfully implemented without consideration of the existing conceptions that teachers have and an awareness of the potentially contradictory uses and functions of assessment in the environment. Change in assessment practice clearly involves changing teacher beliefs, but also requires ensuring environments (e.g., tools and policies) support teachers in deploying the desired changes.

### References

- Barnes, N., Fives, H., & Dacey, C. M. (2015). Teachers' beliefs about assessment. In H. Fives & M. Gregoire Gill (Eds.), *International Handbook of Research on Teacher Beliefs* (pp. 284-300). New York: Routledge.
- Brown, G. T. L. (2008). *Conceptions of assessment: Understanding what assessment means to teachers and students*. New York: Nova Science Publishers.
- Brown, G. T. L. (2012). Teachers' thinking about assessment: Juggling improvement and accountability. *Teacher: the International Education Magazine*, 6(2), 30-35.
- Brown, G. T. L., Hui, S. K. F., Yu, F. W. M., & Kennedy, K. J. (2011). Teachers' conceptions of assessment in Chinese contexts: A tripartite model of accountability, improvement, and irrelevance. *International Journal of Educational Research*, 50(5-6), 307-320. doi: 10.1016/j.ijer.2011.10.003
- Brown, G. T. L., Kennedy, K. J., Fok, P. K., Chan, J. K. S., & Yu, W. M. (2009). Assessment for improvement: Understanding Hong Kong teachers' conceptions and practices of assessment. *Assessment in Education: Principles, Policy and Practice*, 16(3), 347-363. doi: 10.1080/09695940903319737

- 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. doi: 10.1016/j.tate.2010.08.003
- Cizek, G. J. (2001). More unintended consequences of high-stakes testing. *Educational Measurement: Issues and Practice*, 20(4), 19-27.
- Fives, H., & Buehl, M. M. (2012). Spring cleaning for the "messy" construct of teachers' beliefs: What are they? Which have been examined? What can they tell us? In K. R. Harris, S. Graham, & T. Urdan (Eds.), *APA Educational Psychology Handbook: Individual Differences and Cultural and Contextual Factors* (Vol. 2, pp. 471-499). Washington, DC: APA.
- Hamilton, L. S., Stecher, B. M., Marsh, J. A., McCombs, J. S., Robyn, A., Russell, J. L., et al. (2007). Standards-based accountability under no Child Left Behind: Experiences of teachers and administrators in three states. Santa Monica, CA: RAND Education.
- Harris, L. R., & Brown, G. T. L. (2009). The complexity of teachers' conceptions of assessment: tensions between the needs of schools and students. *Assessment in Education: Principles, Policy & Practice*, 16(3), 365-381. doi: 10.1080/09695940903319745
- Hofstede, G. (2007). A European in Asia. *Asian Journal of Social Psychology*, 10(1), 16-21. doi: 10.1111/j.1467-839X.2006.00206.x
- Kennedy, K. J., Chan, J. K. S., & Fok, P. K. (2011). Holding policy-makers to account: exploring 'soft' and 'hard' policy and the implications for curriculum reform. *London Review of Education*, 9(1), 41-54. doi:10.1080/14748460.2011.550433
- Lerner, J. S., & Tetlock, P. E. (1999). Accounting for the effects of accountability. *Psychological Bulletin*, 125(2), 255-275.
- Newton, P. E. (2007). Clarifying the purposes of educational assessment. *Assessment in Education: Principles, Policy & Practice*, 14(2), 149-170.
- Pajares, M. F. (1992). Teachers' beliefs and educational research: Cleaning up a messy construct. *Review of Educational Research*, 62, 307-332.
- Popham, W. J. (2000). *Modern educational measurement: Practical guidelines for educational leaders* (6th ed.). Boston: Allyn & Bacon.
- Ravitch, D. (2013). *Reign of Error: The hoax of the privatization movement and the danger to America's public schools*. New York: A. E. Knopf.
- Rieskamp, J., & Reimer, T. (2007). Ecological rationality. In R. F. Baumeister & K. D. Vohs (Eds.), *Encyclopedia of Social Psychology* (pp. 273-275). Thousand Oaks, CA: Sage.
- Smith, L. F., Hill, M. F., Cowie, B., & Gilmore, A. (2014). Preparing Teachers to Use the Enabling Power of Assessment. In C. M. Wyatt-Smith, V. Klenowski, & P. Colbert (Eds.), *Designing Assessment for Quality Learning* (pp. 303-323). Dordrecht, NL: Springer.
- Stobart, G. (2006). The validity of formative assessment. In J. Gardner (Ed.), *Assessment and learning* (pp. 133-146). London: Sage.