

## **What Might Teachers' Conceptions of Assessment Mean for Validity in High-Stakes, School-Based Assessment?**

Teachers' conceptions of assessment are shaped by and influence their assessment practice. When teachers are the primary architects of high-stakes assessments their conceptions can have powerful repercussions for students and the assessment systems in which they work. This paper provides a qualitative analysis of the conceptions of assessment held by upper secondary teachers in South Australia on the cusp of reforms in senior secondary curriculum, assessment and certification.

Underpinned by Assessment *for* Learning principles the South Australian reforms included an increased reliance on school-based assessment and leading to greater accountability of teachers' assessment practices. Consequently, assessment literacy has moved front and centre in South Australian upper secondary education.

Analysis of the data explores the relationship between pre and in-service learning, teacher accountability in a high-stakes assessment context and the impact of these on educational assessment. A framework that illustrates SACE teachers' cognitive understandings, beliefs and attitudes, approaches and actions in relation to assessment has been developed. The framework is discussed and issues of validity in high-stakes assessment that are raised by the teachers' conceptions are explored.

Cathy Schultz  
Manager, Moderation and Standards  
SACE Board of South Australia  
Adelaide, South Australia  
October 2015  
[cathy.schultz@sa.gov.au](mailto:cathy.schultz@sa.gov.au)

KEY WORDS: conceptions of assessment, senior secondary education, validity, school-based assessment

## Overview of the Study

### Background

The South Australian Certificate of Education (SACE) is the senior secondary certificate that providing pathways to university, vocational education and training and the workplace for South Australian students. The SACE certifies the achievements of students in their final two years of secondary schooling, Year 11 and Year 12. The certificate was introduced in 1992, and subsequently reviewed in 2005-2006, following which a comprehensive agenda of reform in senior secondary curriculum, assessment and certification was undertaken in South Australia.

A key reform was the re-design of the senior secondary assessment system in South Australia to intentionally place greater reliance on teachers' assessment decisions within a high-stakes assessment context. The expanded use of teacher-led assessment in the South Australian Certificate of Education was primarily influenced by two factors. Firstly, a growing momentum within the global education community for using assessment to support and enhance, as well as measure and report, student learning: assessment *for* learning as well as assessment *of* learning. Secondly, a view that assessment validity would be increased "because teacher-led assessment enables a wide range of skills, concepts, processes and understandings to be demonstrated" (Crafter et al., 2006, p.129). Such a view assumes that teachers possess the requisite knowledge and expertise to implement educational reforms in a high-stakes context based on the principles of assessment *for* learning.

### Purpose of the study

The purpose of the study was to explore, from the teachers' perspective, the '*why*' as well as the '*what*' questions of assessment. The research took place during the implementation of the 'new' SACE (2009 – 2012).

### Methodology

Five focus group interviews with thirteen practicing senior secondary teachers<sup>1</sup> were the primary data collection method for the study. Focus group participants were a stratified sample of the broader teaching South Australian population, with reference to the variables of; schooling sector (government, Catholic and independent schools), gender, length and breadth of senior secondary teaching experience, subjects taught and school location (metropolitan vs. rural schools, socio-

---

<sup>1</sup> Focus group participants have been de-identified. Individual teachers are referred to by a unique numeric/ alpha combination (e.g. 1A indicates participant A in focus group 1)

economic quartile of school).

The focus group data were later supplemented by a larger number of responses gathered through a formal evaluation of the implementation of the new SACE (*First Year Evaluation of the SACE - Final Report*, July 2012) and a survey of teachers who participated as SACE moderators (Nov-Dec, 2012). The qualitative research design enabled both the richness and diversity of teachers' individual voices to emerge while, simultaneously, exploring the South Australian senior secondary teaching community's collective conceptions of assessment.

### **Conceptualising 'conceptions'**

Previous researchers have used the term 'conceptions' to encompass "beliefs, meanings, mental images, preferences" (Thompson, 1992, p. 51), a framework for viewing, interpreting and interacting (Marton, 1981) and a framework for understanding and responding (Brown, 2004, Hargreaves, 2005).

The relationship between teachers' beliefs about learning and assessment and their assessment practices has been documented in the assessment literature as a somewhat tense relationship:

In the context of assessment,... although they [teachers] were being asked to take a more systematic approach, they did not always do this: they stuck to their own beliefs and practice in assessing children (Gipps et al., 1999, p.133).

This confusion or ambivalence regarding the relationship between assessment and learning can lead to beliefs about assessment that can be individualistic and/or collectively held and may be an eclectic and often contradictory mix that results in a wide variety of incongruous assessment practice in classrooms (Kahn, 2000). Gavin Brown's work on assessment conceptions identified four diverse and often irreconcilable beliefs about assessment held by New Zealand teachers:

1. Assessment is for improvement of student learning and teacher instruction;
2. Assessment makes students accountable for their learning;
3. Assessment makes teachers and schools accountable for student learning (evaluation);
4. Assessment is irrelevant to the work of teachers (and students) (Brown, 2004)

With reference to centralised policies designed to enhance and improve assessment outcomes for students, such as those envisaged for the new SACE, Brown, recognises that "the success or failure of such policies may hang on the conceptions and meanings that teachers give to these policies" (Brown, 2004, p.301). Similarly, Dixon, Hawe and Parr's assessment *for* learning case study, which explored the 'beliefs practice nexus' in assessment, concluded that "teachers' beliefs are influential in regard to their interpretation, uptake and enactment of assessment reform initiatives" (Dixon et al., 2011,

p.374).

In this study teachers' conceptions of assessment are defined as: an organising framework that encompasses cognitive understandings of assessment, beliefs and attitudes about assessment and learning and approaches and actions taken in the act of assessing.

## **Findings**

The key findings from this study highlight eight cognitive understandings that reflect and influence the South Australian senior secondary teachers' beliefs and attitudes about assessment as well as their approaches and actions taken in the act of assessing. The findings have brought together in a framework (p.5) that illustrates South Australian senior secondary teachers' conceptions of at the time of the study and are discussed in the following section.

### **Assessment is part of the teaching and learning cycle**

South Australian senior secondary teachers understand that assessment is part of the teaching and learning cycle and, as such, is core pedagogical content knowledge (Shulman, 1986, Shulman, 1987). The centrality of assessment to classroom practice leads teachers to view assessment as 'the teacher's job'; a part of their professional identity (Adie, 2013, Harlen, 2006, Nias, 1987, Sadler, 1987, Sadler, 1989), which provides feedback about the effectiveness of teaching practice:

- *1B: It's part of our **job** ..... it's dictated to us.*
- *4A: To find out if I'm doing a good **job** is often a reason why I assess them.*

Although assessment information provides valuable feedback to teachers, it is also seen as a burdensome part of the 'job'. The work involved in 'the teacher's job' of preparing for assessment and making assessment decisions at the *senior* secondary level is seen as stressful and at times resented by, at least some, SACE teachers:

- *3B: I love teaching but the assessment side is hard.*
- *3A: The thing I've found most stressful this year is marking the year 12s.*

The importance of assessment in the teaching and learning cycle is amplified in the upper secondary years of schooling. Participants conveyed an overwhelming sense that assessment at the senior secondary level is more serious by comparison to assessment in the preceding years of secondary schooling, which goes some way towards explaining the sense of burden expressed by teachers in

**TABLE 1: An organising framework of South Australian senior secondary teachers' conceptions of assessment**

Cognitive understandings	Beliefs and attitudes	Approaches and action
Assessment is part of the teaching and learning cycle	Assessment is the teacher's job Assessment is burdensome	Assessment is used to guide teaching and learning Assessment is more serious in senior secondary classes
SACE assessment is high-stakes	Teachers are accountable for the outcomes of senior secondary assessment Teachers need to learn the senior secondary assessment 'game'	The locus of control in assessment primarily rests with teachers Teachers change senior secondary assessment practices in response to curriculum/ system changes
Assessment is an important area of professional expertise for senior secondary teachers	Some teachers have more opportunities, or more actively seek, to develop their expertise in assessment than others  'Luck' plays a role in developing assessment expertise	Some senior secondary teachers assess with more knowledge and confidence than others
Assessment can be used for many purposes (formative, summative and diagnostic)	Assessment enables teachers and students to monitor learning progress and achievement in relation to external requirements and expectations  The purpose of formative assessment is to prepare students for summative assessment	Assessment purposes are blurred in classroom practice Formative assessment replicates summative assessment
Teachers learn to assess by assessing	Experience, particularly moderation and marking experience, develops assessment expertise  Pre-service training provides insufficient preparation in assessment for teaching graduates	Markers and moderators apply their knowledge and understanding of assessment in the classroom  New teachers initially replicate their experiences of assessment (assess as they've been assessed)
Assessment should be fair and equitable	A variety of tasks makes assessment more inclusive and flexible	Assessment programs include a variety of assessment methods  Assessment tasks are differentiated for learners
Assessment should be fit for purpose	Different assessments are appropriate for different subjects	Teachers use subject-specific approaches to assessment
The principles of assessment are absorbed	Teachers need to learn assessment methods and standards but not concepts and principles	Teachers share and seek out assessment practices but not underpinning concepts and principles

this study;

- 1A: *I think the **pressure's** from above with SACE in the higher [years].*
- 1B: (commenting on a colleague's Year 12 assessment practices): *So that teacher had put no effort in whatsoever.... None of that assessment was taken **seriously**.*
- 2A: (talking about the attitudes of senior secondary students' to assessment and learning): *The **serious** aspect of their learning starts to come into play.*

Experienced teachers harness the increased power of assessment in the senior secondary years throughout the teaching and learning cycle. However, neophyte<sup>2</sup> teachers are more inclined to use assessment as a 'bolt on' affair; an action that comes after teaching and learning:

- 3A: *There are times when the sole purpose of assessment is so that I've got a mark to put in my mark book, which is terrible but it's how it is.*
- 3B: *We can be going along doing all this really great stuff in class, doing a novel and really getting into it. And in the back of my mind there's this you've gotta do an assessment at some point.*

Such actions are at odds with the approaches of more experienced colleagues and with contemporary assessment literature that promotes the value of assessment to support learning (ARG, 2002, Black et al., 2007, Black, 1998, Gipps et al., 1999, Harlen and James, 1997, Leighton et al., 2010, Marsh, 2007, William, 2007). Approaches that separate assessment from learning are concerning in terms of the assumptions about teachers' capacities to use assessment information to enhance learning in the new SACE.

### **SACE assessment is high-stakes**

Teachers in this study were cognisant of the high-stakes consequences of SACE assessment for students (e.g. post-school study and employment opportunities), which, in turn, raises the stakes of assessment for SACE teachers. Participants held themselves accountable for students' assessment

---

<sup>2</sup> Those with 1-2 years teaching experience at the time of the study (or recollections of experiences from the early years of an experienced teachers' career)

outcomes and expected their teaching colleagues to do the same;

- 1B: I'm actually **accountable** for my students and if my student's performing at a D level then how am I going to improve that?

- 5A: When I first started teaching [in the lower secondary years] what you did in a way didn't matter all that much. There wasn't that **accountability** factor.

Facilitator: But the more high stakes it is..?

5A: ... you became more and more **accountable**.

- 5B: As teachers we have got that level of **accountability**.....we're expected to deliver successful outcomes for all of the students so I just think that's core business.

However, a negative aspect to these beliefs about teacher accountability in senior secondary assessment is that SACE assessment is viewed as a 'game' for which teachers must learn the rules:

- 5A: Going back 10 or so years ago we went through a thing called group moderation .... And **you learned to play a bit of a game** there because you knew that invariably your marks wouldn't change. So what you did was basically you took your marks up as high as you thought you could.... And then that went out and the statistical moderation came in..... so **you played a slightly different game**..... Now that we're going to the new SACE **we're going to have to learn to play a slightly different game**.

Facilitator: So the quality assurance processes make a difference to what you do in the assessment process?

5A: Yeah, yes absolutely. What I'm trying to do obviously is maximise outcomes for kids. But **you've gotta know the game to play**.

Beliefs and attitudes about 'playing the game' of SACE assessment and maximising assessment outcomes, as opposed to learning outcomes, for students are evidence of the washback of high-stakes assessment (Barnes et al., 2000, Harlen, 2005, Johnson, 2007, Tierney, 2006, Torrance, 2011, Torrance, 2012, Wilmut and Tuson, 2005) on teachers' conceptions of assessment in the South Australian senior secondary context. Many teachers in the study acknowledged that changes in SACE assessment requirements result in changes to classroom assessment practice:

- 2B: Senior school is seen to be driven by an external body which is the SACE driving it as opposed to anything else.... I do quite often look at the assessment section [of the curriculum] first.

- *1A: Towards the end [of secondary school] some people [teachers] focus lots of learning towards the final assessment.*

SACE teachers understand that assessment is part of the teaching and learning cycle and recognize that assessment can be shared by teachers and students (ARG, 2002, Black and Wiliam, 1998, Buhagiar, 2007, Earl, 2003, Pryor and Torrance, 2000). However, the high-stakes nature of SACE assessment, and the resultant accountability that is incumbent upon teachers, appears to limit the partnership between SACE teachers and their students in practice. In focus group teachers' classrooms the locus of control in assessment is almost always with the teacher, with only one participant indicating that student self-reviews or peer assessments are incorporated into their summative SACE assessment practice; raising issues of consequential validity (Johnson, 2007) in SACE teachers' classroom practices.

### **Assessment is an important area of professional expertise for senior secondary teachers**

While assessment is an important area of professional expertise for all teachers, the high-stakes nature of senior secondary assessment increases the importance of assessment literacy for SACE teachers. However, despite the centrality of assessment to professional practice some teachers appear to have more opportunities, or more actively seek, to develop their assessment expertise than others.

Teachers in this study identified advice from mentors/ school leaders, opportunities to embed professional learning in school policy and practice, collegiate planning and discussion of assessment, along with opportunities to engage in a range of experiences as an assessor (classroom, marking, moderation, subject networks) as aspects of a school's assessment culture that made a significant difference to their assessment literacy. In focus group discussions it emerged that such school-based or school-supported opportunities are not available to all SACE teachers. These rich learning opportunities were seen as serendipitous, sporadic and sparse. Participants believe *luck* plays a part in the opportunities teachers have to develop their assessment expertise:

- *1C: ...I can't imagine not having that degree of support [to learn about assessment] among my colleagues and from a coordinator as well.*

*1A: You're very **lucky**.*

*1C: I have been very **lucky** in my teaching career, very **lucky**.*

- *4B: My mentor was my English coordinator, she ... gave me lots and lots of information ... as a*



*mentor she was fabulous.*

*4A: You were very **lucky***

*4D: I got that sort of help later on but not in the first year of teaching.*

- *4B: [Unless] you're **lucky** enough to actually work in a subject that's [socially] moderated.....you really couldn't do all that much [professional learning in assessment].*
- *5A: Yeah – being **lucky** enough to be involved with the setting panel for many years you certainly get a deeper understanding of how questions are set.*

While the teachers in this study unanimously acknowledged the importance of expertise in assessment for senior secondary teachers some exhibited dispositions to 'create their own luck'. According to Ritchhart (2001):

Dispositions concern not only what one can do, one's abilities, but also what one is disposed to do. Thus dispositions address the often-noticed gap between our abilities and our actions (p.144).

The teachers who displayed the highest levels of assessment literacy in the study were also the participants who actively sought out ways to build their capacities in assessment. These teachers regularly engaged in professional reading about assessment, SACE external marking and moderation activities, classroom observations and professional learning networks.

The combination of a school culture that develops assessment expertise and a personal disposition towards assessment literacy means that some SACE teachers approach senior secondary assessment with more knowledge and skill than others. Study participants who taught a wide range of SACE subjects and had been involved in SACE external marking or moderation panels expressed an implicitly deeper understanding of valid, reliable and flexible assessment design. As a consequence, these participants were approaching the new SACE with a confidence and expertise that that would enable a broad range of students in their classes to demonstrate their learning (construct validity) which was likely to benefit their students (consequential validity).

### **Assessment can be used for many purposes**

Focus group discussions demonstrated that South Australian senior secondary teachers understand the power of assessment to measure, support and shape learning. All participants in the study harnessed the power of assessment by routinely using assessment information to monitor students'

progress, make decisions about instruction, and provide feedback to students and report on student learning. Experienced SACE teachers viewed assessment as integral to pedagogy, shunting and weaving between assessment purposes (formative, summative, diagnostic) in their daily practice. Such teachers demonstrate a high degree of inferential validity in their use of assessment as they integrate a range of assessment information to make inferences about student achievement and learning that are “appropriate, meaningful, and useful given the individual or [class they] are dealing with and the context in which [they] are working” (Hubley and Zumbo, 2011, p.228)

However, there was some evidence that participants’ beliefs about teacher responsibility and accountability in senior secondary assessment also lead to “conformative” assessment practices; whereby teachers actively encourage students to act “on the advice they are given in order to meet course objectives” (Torrance, 2012, p.330). For many SACE teachers a near-omnipresent consciousness of the high-stakes consequences of senior secondary assessment leads to formative assessments practices that are in fact “a series of mini-summative assessments” (Harlen, 2006, p.10) that replicate the assessment requirements of SACE subjects. These mini-summative assessments are activities that **teachers** design, **teachers** assess and **teachers** use to and determine the next steps for learning, raising questions about the consequential validity of some SACE assessments.

### **Teachers learn to assess by assessing**

The teachers in this study consistently expressed a view that they entered the teaching profession with insufficient knowledge and skills to enable them to effectively design, interpret and respond to assessment in ways that supported student learning. Consequently, participants unequivocally agreed that experience is the best (and often only) teacher when it comes to assessment. In other words, teachers learn to assess by assessing.

Early career teachers reported that experienced colleagues assumed they had sufficient expertise in assessment and that formalised support to monitor and build assessment literacy was not readily available to them;

- 3B (talking about getting feedback on marking standards): *I was very much aware that I was a young teacher anyway, so I didn't really want them to know that I was really doubting myself...I pretend that I know what I'm doing.*

3A: (affirming 3B's comment about pretending to understand assessment) *And a few times when I actually said to the deputy principal, "I'm still a first year teacher I don't know everything", they just said, "Oh well you're just doing such a good job we forget". And it's*

*like I felt I was doing myself a disservice by pretending all the time.*

In the absence of formal support, those teachers who are most new to senior secondary teaching tend to assess as they were assessed; just as their more experienced colleagues did 35 years before them. For new SACE teachers the “baptism by fire” (4D) at the font of assessment continues.

For many experienced teachers participation in the SACE Board’s external moderation of school-based assessments using social moderation processes (Linn, 1993) filled this vacuum in assessment-focused professional learning. SACE moderation is underpinned by assessment principles and concepts (validity, reliability, equity, comparability, inference, sampling, flexibility), which are articulated in training and reinforced in the moderation process. Thus, throughout the external moderation process teachers “engage in developing a theoretical understanding of the knowledge and skills to be learned” (Hawley and Valli, 1999, p.142).

The external moderation exercise develops and requires evaluative expertise<sup>3</sup> to apply the assessment criteria to student evidence to make decisions about the validity and reliability of the assessment decisions of others. Moderators make inferences about the appropriateness of the assessment decisions of other teachers, on the basis of student work samples from these other teachers’ classes.

The value of SACE moderation as a rich source of professional learning in assessment is recognised by senior secondary teachers; “1A: *To me if anything was to be compulsory of teachers as far as PD goes – it should, every teacher should be made to do a year of moderation at least.*” Teachers in the study who had been SACE moderators unanimously reported that their assessment literacy developed exponentially through participating in external moderation and that they used their enhanced expertise in assessment to improve learning opportunities and outcomes for their students. A survey of 562 SACE moderators in 2012 found that teachers who acted as moderators reported that their understanding of task design (construct validity) and ability to make assessment decisions on the basis of student evidence (inferential validity) was “deeper and richer” (60% of respondents) or “somewhat enhanced” (30% of respondents) by participating in the external moderation process.

---

<sup>3</sup> “Evaluative knowledge can be defined as knowledge of the full set of criteria against which work will be judged and the rules for using these criteria. Evaluative expertise is the ability to make judgements and decisions about work based on the application of multiple criteria” (p.366) DIXON, H. R., HAWE, E. & PARR, J. M. 2011. Enacting Assessment for Learning: the beliefs practice nexus. *Assessment In Education: Principles, Policy & Practice*, 18, 365-379.

## Assessment should be fair and equitable

The assessment principle of equity was well understood and enacted by the teachers in the study. South Australian senior secondary teachers believe that task variety makes assessment programs more inclusive by providing sufficient flexibility to allow all students to demonstrate their learning. The selection of assessment methods by the teachers in this research was usually purposeful. Many participants indicated that they gave particular consideration to ensuring their assessment tasks were engaging and accessible for the cohort of students for whom they were designed. On the surface it appears that the use of a variety of tasks indicates that SACE teachers understand that:

In educational assessments, individual tasks generally are transient and interchangeable and are important only to the extent that they permit meaningful inferences about levels of proficiency and progress within the domain of interest (Masters, 2013, p.39).

However, on closer inspection, the value of providing a variety of tasks seems to be motivated by notions of enhancing student success rather than an understanding of validity across an assessment program:

- *2A: Different students have different passions of learning or different methods of learning and sometimes if we keep our assessment fields too narrow we might miss giving those students the opportunity to be **successful**.....they have to experience **success** in some shape or form.*
- *5B: The flexibility [in SACE assessment tasks] meant that all kids could be very highly **successful** regardless of where they were on the scale..... I genuinely believe that every single kid could be **successful** because you changed or you negotiated the assessment tools to see their learning.*

These sentiments demonstrate these SACE teachers seek to create assessment experiences that impact positively on learners and their learning by ensuring that *all* students' have opportunities to demonstrate their learning. However, there is also an undertone in the teachers' comments which suggests that, in an effort to harness the power of assessment to support learning, SACE teachers may be consciously attending to the assessment principle of equity but ignoring or unconsciously overlooking the principles of validity and reliability. If this were to be the case then questions of comparability arise. Comparability has been defined as:

How well the grades in different courses correspond with the nature, breadth and depth of learning as inferred directly from an integrative and holistic evaluation of all the raw evidence of achievement [student works] (Sadler, 2013, p.6).

Comparability is fundamental in assessment and a necessary pre-condition if high-stakes assessments are to have consequential validity.

### **Assessments should be fit for purpose**

The variety of assessments used by teachers are, in part, a recognition that not only do different *learners* require different assessment but also that “different assessment methods are valid for different kinds of learning” (Masters, 2013, p.38). The teachers in this research held an implicit sense of construct validity; accepting and expecting that the nature of the learning in different subjects would lead to differences in assessment practices between teachers. Consequently, in their assessment practice science teachers used practical experiments, drama and physical education teachers used on-stage/on-field performances, language teachers used conversations, work education teachers used workplace assessments and English teachers used creative writing tasks.

However, while notions of ‘fitness for purpose’ are implied by such assessment practices evidence from the research suggests that the selection and design of tasks is predominantly influenced by the conventions of a subject, which are perpetuated by communities of practice and, for the most part, enshrined in SACE curriculum. Participants did not indicate that their choices of assessment methods were “linked to [a] cognitive model of learning ..... to support the kinds of inferences and decisions that will be based on the assessment results.” (Pellegrino et al., 2001, p.47). Consequently, participants’ assessment practices were subject-specific on the basis of prescription or tradition rather than being the result of an explicit consideration of:

The method’s capacity to provide information about where students are in their learning within the domain of interest – in other words, its construct validity or fitness for purpose”. (Masters, 2013, p.38).

The lack of a cognitive model of learning means that opportunities for valid assessment practices to cross subject boundaries are rarely exploited by teachers in the senior secondary years. Therefore, while SACE teachers use a variety of assessment practices these practices are “usually balkanised into subject-based departments, each with its distinctive culture” (Timperley et al., 2007, p.208).

### **The principles of assessment are absorbed**

Despite the consensus among participants that pre-service training had not adequately prepared them for classroom assessment, and that opportunities to learn about assessment are serendipitous, there seemed to be an attitude that assessment concepts and principles are absorbed as teachers develop

their assessment practices. To an extent this is true, with some implicit understandings of assessment principles revealed during the focus group discussions. However, in the discussions it appeared that the professional language of assessment and some key assessment concepts, such as inference, comparability, sampling, and random and systematic errors were not as readily absorbed.

Teachers believe they can assess effectively once they have developed an implicit understanding of what constitutes a 'good' task and how to mark to the 'correct' standard. While 'good task' and 'correct standard' can be seen as implicit understandings of the concepts of validity and reliability, study participants expressed a view that many teachers are expeditious in their professional learning about assessment; seeking out practical skills and resources ahead of theoretical understanding:

- *4A: They [teachers] don't wanna know the philosophies behind it. They just wanna know, "What do I have to do and how can that improve the outcome for the kids?" and that's it.*

*4B: I've presented hundreds of sessions at different workshops ..... and the first thing I get is, "What worksheet are you going to give the kids tomorrow?"..... It's not about anything other than, "What practical resource are you going to give me to make my life better?"*

- *5B: I personally haven't done anything [professional learning] other than what has been done through ... the SACE Board. And I s'pose that's just a time thing and prioritising ... I just work within the [assessment] model that we've got at that particular time.*

Such selective and pragmatic approaches to professional learning are efficient but can also lead to a myopic understanding of assessment; that is an understanding of 'what works' in *my* subject, for *my* students, for the assessment system *I* work in. An absence of a 'big picture' theoretical understanding of assessment can lead to gaps in assessment literacy that make it difficult for teachers to transfer their practices from one context. As a result they may feel, or indeed be, vulnerable when the context in which they are working changes; for example, when moving to a new school, teaching a new subject or working with a new assessment model.

Such 'gaps' in assessment literacy appear to be evident in responses from the broader teaching population highlighted in the *First Year Evaluation of the SACE – Final Report*, a report prepared by an independent evaluation panel after the first cohort of students completed Year 12 under the new SACE in 2012. Respondents expressed a number of concerns in relation to assessment that appear to be tied to misunderstandings, or partial understandings of assessment and measurement theory as outlined in the table below.

**Table 2: SACE teachers concerns about SACE assessment in 2012**

Concern	Assessment principles/concepts
Task design	<ul style="list-style-type: none"><li>• Validity</li><li>• Comparability</li></ul>
Performance standards (and marks vs. grades)	<ul style="list-style-type: none"><li>• Reliability</li><li>• Comparability</li></ul>
Social moderation vs. statistical moderation	<ul style="list-style-type: none"><li>• Equity</li><li>• Measurement error (and random vs. systematic errors)</li><li>• Sampling</li><li>• Inference</li></ul>

### **Implications for validity**

Messick views validity as a unified but multi-faceted concept that requires “an integrated evaluative judgement” (Messick, 1989b). Construct validity, inferential validity and consequential validity, three key facets of Messick’s concept, have been brought to the fore in this study of South Australian teachers’ conceptions of assessment; namely. His definition takes into account, not only the extent to which an assessment measures what it has been designed to measure, but also the conclusions or inferences made on the basis of the assessment and the consequences of those conclusions.

The teachers in this study demonstrated an implicit understanding that construct validity means different types of learning require different types of assessment. However, notions of ‘engagement’ and student ‘success’ appear to take precedence over construct validity in classroom practice. Tasks are developed, negotiated and adapted to cater for students’ interests and abilities but consideration of whether the tasks administered actually sample the domain (or construct) of interest was notably absent in focus group discussion. It is certainly possible that these school-based assessment tasks do indeed allow valid inferences to be made about student learning in relation to the learning domain (Cronbach, 1971). But, it is equally possible that construct under-representation or construct irrelevant variance (Messick, 1995, Messick, 1989a) are evident if teacher-designed tasks value equity and access at the expense of construct validity.

As Kyriakides suggests, the teachers in this study “as end-users of tests, contribute a distinctive perspective on ... inferential validity” (Kyriakides, 2004, p.144). The view of SACE assessment as a game, with rules that must be learned, leads some teachers to make interpretations, decisions and actions (Moss et al., 2006, p.111) using inferences about potential summative achievement at the expense of inferences about learning progression (Masters, 2013). In seeking to “maximise outcomes for kids” (5A) teacher focus is drawn towards the assessment system in which they are working, rather than the validity of the inferences they are making about student learning on the basis of the assessments they conduct.

It has been argued that “you must evaluate the intended consequences and unintended side effects of measurement when validating the inferences and uses made from tests and measures” (Hublely and Zumbo, 2011, p.220). The consequential validity or ‘washback effect’ of high-stakes assessment is well-documented globally and not unique to the conceptions of South Australian teachers. For SACE teachers the washback of assessment is often positive with the ‘seriousness’ of senior secondary assessment spurring them to learn more about assessment through professional learning communities and participation in system-wide marking and moderation processes (an intended consequence). However, the accountability of SACE assessment also leads to professional pragmatism that may limit innovation and experimentation in senior secondary school-based assessment. Teachers reported that they, and their colleagues, seek to develop a bag of tried and true assessment ‘tricks’. This ‘bag of tricks’ typically contains assessment tasks with which teachers have observed, or experienced, professional success and is an unintended consequence that raises questions of consequential validity for high-stakes, school-based assessment systems such as the South Australian Certificate of Education .

## **Conclusion**

Teachers’ conceptions of assessment have a direct impact on teaching and learning in classrooms and the opportunities afforded for students to demonstrate what they know, understand and can do. The extent to which teachers’ individual and collective conceptions influence assessment validity is worthy of further exploration. Ultimately, the validity of high-stakes school-based assessments depends not on the curriculum writers or assessment experts but on teachers’ cognitive understandings, beliefs and attitudes, actions and approaches taken in the act of assessing.



## References

- ADIE, L. 2013. The development of teacher assessment identity through participation in online moderation. *Assessment In Education: Principles, Policy & Practice*, 20, 91-106.
- ARG. 2002. *Assessment for Learning, 10 principles* [Online]. Available: [www.assessment-reform-group.org](http://www.assessment-reform-group.org) 2007].
- BARNES, M., CLARKE, D. & STEPHENS, M. 2000. Assessment: the engine of systemic curricular reform? *Journal of Curriculum Studies*, 32, 623-650.
- BLACK, P., HARRISON, C., HODGEN, J., MARSHALL, B. & SERRET, N. 2007. Riding the interface: An exploration of the issues that beset teachers as they strive for assessment systems. *British Educational Research Journal*.
- BLACK, P. & WILIAM, D. 1998. Assessment and classroom learning. *Assessment in Education*, 5, 7-74.
- BLACK, P. J. 1998. *Testing: friend or foe? The theory and practice of assessment and testing*, London, Falmer Press.
- BROWN, G. T. L. 2004. Teachers' conceptions of assessment: implications for policy and professional development *Assessment in Education: Principles, Policy & Practice*, 11, 301-318.
- BUHAGIAR, M. A. 2007. Classroom assessment within the alternative assessment paradigm: revisiting the territory. *The Curriculum Journal*, Vol. 18, 39 - 56.
- CRONBACH, L. J. 1971. Test Validation. In: THORNDIKE, R. L. (ed.) *Educational Measurement*. 2nd. ed. Washington, DC: American Council on Education.
- DIXON, H. R., HAWE, E. & PARR, J. M. 2011. Enacting Assessment for Learning: the beliefs practice nexus. *Assessment In Education: Principles, Policy & Practice*, 18, 365-379.
- EARL, L. 2003. *Assessment as Learning: Using Classroom Learning to Maximize Student Success*, California, Corwin Press.
- GIPPS, C., MCCALLUM, B. & BROWN, M. 1999. Primary teachers' beliefs about teaching and learning. *Curriculum Journal*, 10, 123 - 134.
- HARGREAVES, E. 2005. 'Assessment for learning? Thinking outside the (black) box'. *Cambridge Journal of Education*, 35, 213 - 224.
- HARLEN, W. 2005. Teachers' summative practices and assessment for learning - tensions and synergies. *The Curriculum Journal*, 16, 207-223.
- HARLEN, W. 2006. The role of teachers in the assessment of learning. In: GROUP, A. R. (ed.). Newcastle Document Services.
- HARLEN, W. & JAMES, M. 1997. Assessment and learning: differences and relationships between formative and summative assessment. *Assessment in Education: Principles, Policy and Practice*, 4, 365-379.
- HAWLEY, W. D. & VALLI, L. 1999. The essentials of effective professional development: A new consensus In: DARLING-HAMMOND, L. & SYKES, G. (eds.) *Teaching as the Learning Profession: Handbook of Policy and Practice*. San Francisco, CA: Jossey-Bass.
- HUBLEY, A. M. & ZUMBO, B. D. 2011. Validity and the Consequences of Test Interpretation and Use. *Social Indicators Research*, 103, 219-230.

- JOHNSON, N. 2007. A consideration of assessment validity in relation to classroom practice. *IAEA Annual Conference*. Baku Azerbaijan.
- KAHN, E. A. 2000. A case study of assessment in a grade 10 English course. *The Journal of Educational Research*, 93, 276-287.
- KYRIAKIDES, L. 2004. Investigating validity from teachers' perspectives through their engagement in large-scale assessment: The Emergent Literacy Baseline Assessment project. *Assessment in Education: Principles, Policy & Practice*, 11, 143-165.
- LEIGHTON, J. P., GOKIERT, R. J., COR, M. K. & HEFFERNAN, C. 2010. Teacher beliefs about the cognitive diagnostic information of classroom- versus large-scale tests: implications for assessment literacy. *Assessment In Education: Principles, Policy & Practice*, 17, 7-21.
- LINN, R. L. 1993. Linking results of distinct assessments. *Applied Measurement in Education*, 6, 83-102.
- MARSH, C. J. 2007. A critical analysis of the use of formative assessment in schools. *Educational Research Policy and Practice*, 6, 25-29.
- MARTON, F. 1981. Phenomenography - describing conceptions of the world around us. *Instructional Science*, 10, 177-200.
- MASTERS, G. N. 2013. Reforming Educational Assessment: Imperatives, principles and challenges. *Australian Education Review*, 57.
- MESSICK, S. 1989a. Meaning and values in test validation: The science and ethics of assessment. *Educational Researcher*, 18, 5-11.
- MESSICK, S. 1989b. Validity. In: LINN, R. L. (ed.) *Educational Measurement*. 3rd ed. New York: MacMillan.
- MESSICK, S. 1995. Validity of Psychological Assessment. *American Psychologist*, 50, 741-749.
- MOSS, P., GIRARD, B. & HANIFORD, L. 2006. Validity in educational assessment. *Review of Research in Education*, 30, 109-162.
- NIAS, J. 1987. Teaching and the self. *Cambridge Journal of Education*, 178-184.
- PELLEGRINO, J. W., CHUDOWSKY, N. & GLASER, R. 2001. *Knowing what students know: The science and design of educational assessment*, Washington, DC, National Academy Press.
- PRYOR, J. & TORRANCE, H. 2000. Questioning the Three Bears. In: FILER, A. (ed.) *Assessment: Social Practice and Social Product*. London: RoutledgeFalmer.
- RITCHHART, R. 2001. From IQ to IC: A dispositional view of intelligence. *Roeper Review*, 23, 143-150.
- SADLER, D. R. 1987. Specifying and Promulgating Achievement Standards. *Oxford Review of Education*, 13, 191-209.
- SADLER, D. R. 1989. Formative assessment and the design of instructional systems. *Instructional Science Education*, 18, 119-144.
- SADLER, D. R. 2013. Assuring academic achievement standards: from moderation to calibration. *Assessment in Education: Principles, Policy & Practice*, 20, 5-19.
- SHULMAN, L. S. 1986. Those who understand: Knowledge growth in teaching. *Educational Researcher*, 15, 4-14.

- SHULMAN, L. S. 1987. Knowledge and teaching: Foundations of the new reform. *Harvard Educational Review*, 57, 1-22.
- THOMPSON, A. G. 1992. Teachers' beliefs and conceptions: A synthesis of the research. In: GROUWS, D. A. (ed.) *Handbook of research on mathematics teaching and learning*. New York: Macmillan.
- TIERNEY, R. D. 2006. Changing practices: influences on classroom assessment. *Assessment in Education*, 13, 239-264.
- TIMPERLEY, H., WILSON, A., BARRAR, H. & FUNG, I. 2007. Teacher professional learning and development: Best evidence synthesis iteration [BES]. Wellington: Ministry of Education.
- TORRANCE, H. 2011. Using Assessment to Drive the Reform of Schooling: Time to Stop Pursuing the Chimera? *British Journal of Educational Studies*, 59, 459-485.
- TORRANCE, H. 2012. Formative assessment at the crossroads: conformance, deformative and transformative assessment. *Oxford Review of Education*, 38, 323-342.
- WILIAM, D. 2007. Sensitivity to instruction: the missing ingredient in large-scale assessment systems? *33rd. IAEA conference*. Baku, Azerbaijan.
- WILMUT, J. & TUSON, J. 2005. Statistical Moderation of Teacher Assessments. Qualifications and Curriculum Authority, UK.