

Formative assessment as mediation

MARK DE VOS

Rhodes University

DINA ZOË BELLUIGI

Rhodes University

Whilst principles of validity, reliability and fairness should be central concerns for the assessment of student learning in higher education, simplistic notions of ‘transparency’ and ‘explicitness’ in terms of assessment criteria should be critiqued more rigorously. This article examines the inherent tensions resulting from CRA’s links to both behaviourism and constructivism and argues that more nuance and interpretation is required if the assessor is to engage his/her students with criterion-based assessment from a constructivist paradigm. One way to negotiate the tensions between different assessment ideologies and approaches meaningfully is to construe assessment as ‘mediation’. This article presents an example assessment rubric informed by John Biggs’ (1999) SOLO Taxonomy.

Keywords: Criterion Referenced Assessment, CRA, mediation, arbitration, formative assessment, Bloom’s Taxonomy

Introduction

South African higher education policy has been moving towards the adoption of criterion-referenced assessment (CRA) in aid of Outcomes-Based Education (OBE), in conjunction with an increased focus on constructivist approaches (Boughey, 2004).¹ CRA models’ requirements for explicit outcomes and assessment criteria are perceived as central both to the commoditisation of education and to ensuring accountability to the State in service of societal transformation (Morrow, 2007; Singh, 2001). In such a context, critique of the notion of transparency in assessment criteria has been minimal, resulting in widespread compliance (for critiques of transparency in other contexts, see Strathern, 2000; Knight, 2001; Parker, 2003).

Tensions arising from different conceptions of CRA

Criterion-referenced assessment (CRA) is often polemically contrasted with norm-referenced assessment (NRA), which was the conventional form of assessment in the majority of South African HE institutions until recently. CRA is offered as the more sound alternative, where the performance or achievement of the student is referenced against standards which are determined and made explicit before the assessment event. However, these distinctions are somewhat artificial, with some arguing that “today’s over-reliance on explicit knowledge could perhaps be as naive as the over-reliance on tacit knowledge had been in the past for the communication of assessment criteria and standards” (O’Donovan, Price & Rust, 2004: 327).

In accordance with this artificial distinction, assumptions behind CRA emphasise the role which assessment plays in modifying student behaviours. Those informed by the competency movement would contend that assessment should be graded against criteria ranging from complete lack of competence through to complete and perfect competence, in the spirit of Glaser (1963). Although this is often practised by educators informed by constructivism, CRA in such applications can be perceived to be rooted in behaviourism. This entails that assessment criteria must be specified in advance and should ideally specify completely the ranges of behaviour that a student should be able to demonstrate in the assessment event. Knowledge of learning

¹ While we acknowledge that CRA is not necessarily linked to OBE, in the South African context, they often go hand in hand. Thus, the remainder of this article will take it for granted that the South African implementation of OBE is strongly linked to CRA.

outcomes and criteria articulate to students the specific intentions of the course. The aim is to describe to students what they should know, understand and be able to do on completing the course.

Against this instrumentalist perspective are conceptions where the focus of learning outcomes is not on objective-based teaching (Knight, 2001, among others). There is a difference between the behaviourist slant of shaping curricula around what a student should be able to 'do' and the constructivist curriculum as a "set of purposeful, intended experiences" (Knight, 2001: 369). Constructivist approaches are geared towards emphasising the value of learning and teaching processes. Constructivist approaches make the educator more reflexive in his/her teaching and more sensitive to his/her students' needs and diversity. Behaviourist conceptions emphasise the intended result of studying, while more constructivist conceptions emphasise the process of learning during the studies.

These two conceptions betray an ideological tension in the use of CRA. CRA is used, on the one hand, in a behaviourist way to assess performance and, on the other, in a constructivist way to co-construct understandings with learners where students actively interpret this information.

The problem of transparency and objectivity

A point of conflict crucial to this article is how these tensions play out in relation to the supposed transparency and objectivity of the assessor. Behaviourist-based approaches require positive and negative reinforcement of student behaviours. This means that they require explicit formulation of criteria in advance; if there are no explicit criteria, there can be no objective standard on which to base reinforcement. In addition, the clarity of such criteria has been considered useful for the constructive alignment of assessment with outcomes (Popham, 1993; Biggs, 1999). However, clear criteria do not suffice to ensure uniform interpretation (Millman, 1994: 19; Knight & Yorke, 2008: 178). It is often not acknowledged that "there is a degree to which criteria cannot be unambiguously specified but are subject to social processes by which meanings are contested and constructed" (Knight, 2001: 20). For this reason many proponents of CRA actively reject high levels of description or precision (O'Donovan, Price & Rust, 2000; Elander, Harrington, Norton, Robinson & Reddy, 2006; Sadler, 2005; Hammer, 2007).

By contrast, constructivism requires criteria to be negotiated in context rather than set in stone in advance. In this way, the gap between 'expert' and student is not perceived as defective but rather as "normal learning awaiting further development" (Francis & Hallam, 2000: 295). Consequently, there is a tension between the behaviourist approaches' explicitness and mechanical ease of application and the constructivist approaches' subjective and negotiated, discursive construction of assessment positions.

Since many measurement theorists seem to focus more on theoretical questions, large-scale test development, validity and reliability than on classroom interventions (Smith, 2003), it is not surprising that this aspect of detailed CRA has to a large extent been overlooked. Only relatively recently have educationalists such as Shay (2004; 2005; 2008) argued that assessment should be recognised as a socially situated interpretative practice, what Killen (2003: 1) calls "an integrated evaluative judgement". Similarly, Knight (2001) contends that assessment is local practice, because the meanings of assessment activities are bound up with the particular circumstances of their production. This requires some reconsideration of the argument for precise transparency or objectivity in order to arrive at a more nuanced alternative.

Possible solutions to this quandary

Assessors who use CRA find themselves in some kind of quandary. On the one hand, transparency, explicitness and precision are the core justification for CRA; on the other, this drive can make CRA standards unattainable in theory and unusable in practice. Despite this paradox, the lecturer is not absolved of his/her responsibility to providing students with epistemological access through insight into assessment practices.

One possible way to negotiate this quagmire is to use 'indicators' instead of criteria (Knight, 2001). This requires replacing highly detailed, outcome-referenced assessment with more general descriptors that highlight the cognitive skills or the 'cognitive essence' (Popham, 1993: 13) which students need to practise. This idea is developed by Sizmur & Sainsbury (1997) who argue that the fact that criterion-

referenced behaviours have some kind of educational value implies that a theoretically prior notion exists which would allow the behaviours specified in outcomes to be mapped to more general cognitive skills.

Another approach to the problem is to move away from precise descriptors towards privileging the relationship between the lecturer and the student. Hussey & Smith (2002: 359) propose an ‘articulated curriculum’ which moves away from the focus on assessment criteria to the teaching-learning relationship. Through a diverse range of assessment practices, including scaffolded tasks and self- and peer-assessment, lecturers and students can share in the responsibility of attempting to create ‘shared understandings’ of the meanings of criteria (Niven 2009). However, the assumption that shared understandings can be achieved in the first place is problematic because the range of understandings possible may be infinite. Moreover, there is always the philosophical question of how to determine whether an understanding is truly ‘shared’ or not.

The politics of mediation

It has been argued that many current applications of CRA experience a number of problems, including exhaustively specified descriptors and a denial of the interpretative nature of the process of assessment, perhaps because of a behaviourist grounding that is incompatible with teaching and learning conceptions informed by constructivism. Some of these problems can be overcome if assessment is reconceptualised as mediation (as opposed to arbitration).

Outside of the HE context, mediation has been defined as a process of conflict settlement where a putatively neutral mediator facilitates the communication between two parties but does not impose a specific agreement (Silbey & Merry, 1986). This involves the creating of a resolution, starting from the prior understandings of the participants and ultimately changing some of those understandings as a result of the process of mediation. This approach is in many ways suited to the teacher-assessor’s mediation of the conflict inherent in transformative learning, which necessitates movements between equilibrium and dissonance, between the students’ prior knowledge and desires and that of the professional community of practice into which they are being inducted.

The differences between the notion of mediation and that of arbitration or adjudication as used in legal systems should be noted. Arbitration involves a third party to act as judge. Unlike a mediator, an arbitrator makes no attempt to appear neutral but actively takes a side. Granted institutional power, an arbitrator does not have to obtain the confidence of both sides and will rule in favour of one of the participants rather than coming to a new, negotiated settlement. Enforcement of the result is often coercive in the sense that it is upheld by the law and/or the police services.

Mediators have an inherently contradictory task (Jacobs, 2002). They need to (a) appear neutral even though they obviously have vested interests; (b) obtain the confidence of both sides in the process and final outcome of the mediation; (c) move both sides towards new negotiated common ground, (d) without the use of overt coercion. These demands often place the mediator in situations with ‘paradoxical expectations’ (Silbey & Merry, 1986: 7). Neutrality, or rather perceived neutrality, and its construction are powerful tools used to negotiate these tensions (Jacobs, 2002) and legitimate the mediation process (Field, 2002; Maiese, 2005; Douglas & Field, 2006). However, the mediator is, in fact, not neutral at all, but is an active participant in the process and wields considerable power (Silbey & Merry 1986; Jacobs, 2002; Boule, 2005 *inter alia*).

For instance, mediators may control the mediation process by (a) presenting themselves as having institutional support, being experts in the discipline, *etc*; (b) controlling the mediation process and the communication that occurs within it; (c) selecting and framing the issues under discussion and constructing a new account of the conflict and its understandings, and (d) ‘activating commitment’ by presenting the outcome of mediation in ideological terms (Silbey & Merry, 1986). This allows the mediator to move the discussion towards resolution without having to exert overt power and thereby threaten the face of the participants.

Assessment as mediation

There are distinct parallels between mediators and teaching professionals in the context of HE. Assessment and mediation share the same basic rationale. Mediation is about resolving conflicts of various kinds. Assessment is, at heart, about conflicting understandings of various issues. For instance, a student expresses his/her understandings in an assignment, which may be similar or different to the 'canon'. Gradually, as a result of sustained engagement with formative assessment practices, the student develops new and deeper understandings (Black & Wiliam, 1998). Ultimately, at advanced postgraduate level, the student may gain insights that challenge or change the understandings of the professional community (indeed, this is often expected at PhD level). Thus the professional community may ultimately adjust its own theory, approaches or practice in response to (admittedly advanced) learning on the part of the student.

Like a mediator, an assessor facilitates interaction between two other parties, the student and the community of practice – although this is not obvious at first sight as the only two parties immediately evident in the classroom are the assessor and the student. Constructivist approaches to learning tend to emphasise the assessor-student relationship as a two-way interaction, often characterised as 'negotiation' or 'facilitation' between two parties. A negotiator must come to terms with a situation where both parties are more or less equals; they each have something to bargain with. It appears that this does not reflect the socio-political dynamics of the classroom: it is not the case that students negotiate outcomes with the assessor directly. Rather, the assessor has institutional power over the students, but in the interests of mediating learning outcomes, the assessor may choose to mask overt expression of that power. For this reason, we prefer to distinguish assessment as mediation from approaches which conceptualise assessment as negotiation.

What the constructivist conception of assessment as negotiation often omits is that the involvement of an additional third party in the learning process is necessary. This third party is the knowledge community into which students are being inducted (Shay, 2004; 2005; Price, 2005) and which often has institutionalised gatekeepers (such as the bar exams for legal practitioners, articles for accountants, professional licenses for medical practitioners, *etc.*). Assessors act as mediators between that community and their students as novices in this field.

The substantive difference between mediation of assessment in HE and that in other contexts is that in the former, one party is not physically present in the classroom. However, the professional community's presence is reified in various ways in the classroom: in textbooks and seminal texts, in examination memos and examples of best practice, in moderators and external examiners, in the 'expert' knowledge and role of the assessor, *etc.* Learning and assessment thus become at least a three-way process – a multilogue – where an assessor mediates between the student and a broader professional community.

Like mediators, the assessor must maintain the confidence of both the student and the professional communities of practice in educational outcomes (Douglas & Field, 2006: 199; Niven 2009: 281). Professional communities must believe that the assessor/educational system is producing high-quality students. Students must believe that the assessor has their best interests at heart in teaching relevant skills and material that will, in turn, make the student acceptable to the professional community. Although this is often not acknowledged, it is a reason why in practice assessors strive so hard to ensure reliability, avoid the possibility of plagiarism, utilise external examiners and moderators, *etc.*

Similarly, mediators place a heavy focus on maintaining the illusion of perceived neutrality. This concern for neutrality seems not to be present to the same degree in HE assessment literature, where more emphasis is placed on the principles of reliability and validity of assessment (cf. Killen, 2003). However, assessors are definitely not neutral insofar as they have a vested interest in facilitating students' understandings of new material and in providing bridges between students' prior knowledge and new domains. Furthermore, assessors are typically experts in their fields and may even be practitioners of their disciplines. Thus, the assessor often has a strong vested interest in the community of practice. S/he must aid students in adjusting their position (with regard to knowledge and skills) to ensure agreement between the students' position and the requirements of the professional community. Formative assessment should be inherent in this approach.

However, engagement in an issue can lead to ethical dilemmas since attempting neutrality can disrupt what some perceive as their “ethical duty to ensure just outcomes” (Douglas & Field, 2006: 186). Many mediators argue for an advocacy role (Jacobs, 2002; Field, 2002), where the mediator actively advances the cause of the weaker party. Assessors, who have a sense of education as emancipatory or empowering, or where curriculum is conceived as praxis, might find such potential compelling.

These parallels strongly suggest that constructivist assessment practice can be theorised in terms of Mediation Theory. More research is needed on other models of mediation in relation to other models of assessment. For instance, summative assessment tends to emphasise the arbitrational role of assessors. The fact that this may be the case does not undermine the argument of this article; all it shows is that there are differences between formative and summative assessment practices. However, it would be interesting for future research to explore this idea further and determine whether assessment-as-arbitration can be constructed as a counterpoint to the proposal in this article that formative, constructivist assessment practices can be regarded as mediative in character.

Mediating standards

The previous sections argued broadly (i) for a constructivist, mediative context for assessment and (ii) for descriptors that focus on underlying, general intellectual constructs rather than mechanistic, atomised specifications for assessment criteria. Given the well-documented problems in attempting to define precise specifications, we propose that we do not even try. Rather, drawing on the insights of theorists arguing for an appropriate grounding of standards, each standard is couched in cognitive processes and intellectual abilities which capture the ‘intellectual essence’ (Popham, 1993) of a particular activity. It is thus important to define the standards with this ‘intellectual essence’ without falling into the trap of an infinite regression of explicitness.

One way of engaging with standards for assessment is by using educational taxonomies, such as Benjamin Bloom’s taxonomy (Bloom, 1956; Krathwohl, 2002), Perry’s Scheme of Intellectual and Ethical Development (Perry, 1970), and John Biggs’ taxonomy (1999; 2003), *etc.* According to such taxonomies, cognitive capacities develop from simple to complex (higher order) abilities or skills. For example, the recall of facts may be subsumed by comprehension, extrapolation, application, *etc.*

Whatever taxonomy is chosen, its standards should allow implicatures (Grice, 1975) about the *manner* in which students achieve the outcomes (Killen, 2003: 10).² Note that this is different to behaviouristically influenced CRA practices that stress what a student can in fact *do* but place less emphasis on the *manner* in which it is achieved. Thus, Bloom’s taxonomy is a system for classifying learning outcomes, not necessarily a system for classifying *how* students achieve them.

To illustrate this difference, consider a hypothetical outcome of an English writing course: *Write a haiku*. From the perspective of Bloom’s taxonomy (1956), such an outcome can be classified as *Synthesis*, or *Creating* (Krathwohl, 2002; Knight, 2001). There are clearly different levels at which the writing of a haiku could be achieved: a student could compose a text that is formally a haiku (one which has the requisite syllable structure) without necessarily obtaining the emotional and/or succinctness qualities that make the form poetic. Alternatively, a student could write a haiku which makes sense, but which does not constitute a ‘great’ haiku, *etc.* Clearly, all of these responses are part of the *Create* category in Krathwohl’s terms, but they fall along a range of achievement. Furthermore, the evaluation of each of these haikus requires an interpretive act on the part of the assessor (Knight, 2001; Shay, 2004; Shay & Jawitz, 2005) that mediates between the student’s haiku and the extended community of haiku critics and poets, and a canon of literature.

2 Although this article may not be the appropriate context to discuss the types of ‘meanings’ derived from taxonomies’ standards, in the interests of precision, we have opted to use the technical term ‘implicature’ rather than remaining vague about what we mean. By ‘implicature’ we mean the conversational implicatures identified by the philosopher, Paul Grice (1975). Implicatures are contextually derived, non-entailed meanings generated as a result of the interaction between the contextual expectations of the decoder and the apparent entailments of the message in accordance with the Cooperative Principle.

We adopt Biggs’ taxonomy (1999) below, because it usefully allows each level of achievement – *Prestructural, Unistructural, Multistructural, Relational* and *Extended Abstract* – to be characterised in an interpretive manner.³ Biggs’ taxonomy is not necessarily a measure of (behaviourist) acts (typically characterised by verbs (Spady, 1994)) but rather of manners of achievement (typically characterised by adverbs and adjectives) in which acts are implemented. Thus, it is possible to write a haiku in a unistructural way, a multistructural way, a relational way, *etc.* As the student progresses through postgraduate study, the *Relational* and *Extended Abstract* categories become the defining characteristics of research.

Table 1: Proposal for a mediated CRA grid

Outcomes The student fulfils the outcome in a ...	Prestructural manner	Unistructural manner	Multistructural manner	Relational manner	Extended abstract manner	Challenging manner
Write a Haiku						
Explain the relationship between Haiku and other poetic forms						

However, Biggs’ taxonomy as it stands needs to be developed for a mediative approach to assessment. In its current form it does not necessarily provide the space for a student to exceed the expectations imposed by the taxonomy. To create a space for this kind of excellence at earlier levels, we have added another category: *Challenging*. This gives the student licence to achieve a particular outcome in a manner which the assessor and/or the professional community to reassess their own understandings of the discipline and/or assessment. This is partly to overcome concerns that while transparency and objectivity are clearly desirable within assessment practices, an overemphasis on them risks stultifying the learning process, marginalising exceptional qualities and consequently dampening leaps of thought and practice (Gordon, 2004). Such leaps are the mark of the most insightful students. Some students may surpass expectations, answer or problematise questions in new and exciting ways, and actively seek new knowledge. In a highly descriptive, behaviourist, CRA approach there is no room to accommodate these exceptional responses. An effective CRA method should not only accommodate these responses, but actively promote them. After all, a truly successful educator is one whose students challenge his/her understandings of the subject matter.

At this point, it is worth mentioning that Reed, Granville, Janks, Makoe, Stein, Van Zyl & Samuel (2002) have implemented a similar rubric using the Biggs taxonomy. The proposal presented in this instance differs from theirs in many respects. First, Reed *et al* (2002) develop a rubric that is a form of summative assessment. The mediative approach developed here is entirely focused on formative assessment.

Secondly, Reed *et al* (2002) root their proposal in the need for inter-marker reliability of grading (i.e. measurement). The mediative approach regards measurement as a different issue entirely that is ultimately

3 We are aware that Biggs argues in favour of explicit standards and that in this article we argue against them. However, nothing about Biggs’ taxonomy necessarily entails explicitness. Furthermore, our choosing Biggs’ taxonomy does not mean that this is the only possible set of standards that could be used; it is merely one that fits the criteria we have outlined.

bound up with summative assessment. The calculation of the final mark can be done separately and is in a sense a completely different operation.⁴

Thirdly, Reed *et al* (2002) provide a set of explicit criteria. They acknowledge the inadequacies of this in dealing with context-dependent considerations (e.g. their discussion of ‘voice’) and point out that their proposed rubric is inadequate to deal with them. Under the mediative approach, Biggs’ categories are themselves the specification of the standards. The assessor must ultimately make a fine-grained analysis of whether the student’s work constitutes a prestructural, unistructural, multistructural, relational, *etc.* response to the requirements of the task. In doing so, the assessor must make implicatures or validity judgements (Killen, 2003; Shay & Jawitz, 2005) from the student’s work to the assessment category and must explain why this assessment is justified. An assessment of a particular outcome could thus take the general form, “*For outcome 1, this work appears to be multistructural because ...*”. As Black & Wiliam (1998: 9) put it: “Feedback to any pupil should be about the particular qualities of his or her work with advice on what he or she can do to improve ...”. This analytical process is essentially context-dependent and mediative in character.

It may be argued that this interpretive process is inimical to CRA on the grounds that it is unclear and inexplicit. We would counter that the skill of interpretation is a desirable and authentic (Wiggins, 1990) outcome of HE and so, by acknowledging and incorporating the need for interpretation into assessment procedures, educators are in fact demonstrating critical alignment between outcomes and assessment.

The final difference between this proposal and that of Reed *et al* (2002) is that in developing explicit criteria, Reed *et al* do not necessarily leave open the possibility of exceeding them. In the mediative approach, the possibility that students could challenge the conceptions of the assessor and/or the professional community is left open. This is codified in the *Challenging* category we have included.

Conclusion

This article argued that a fully explicit set of CRA criteria is problematic in theory and unwieldy in practice. We have drawn on Mediation Theory to design what we argue is an improved CRA system. We believe that this approach to formative assessment can more directly assist with epistemological access to the knowledge construction of communities of practice outside of HE, without paying uncritical lip service to notions of explicitness and transparency.⁵

References

- Biggs J 1999. What the student does: Teaching for enhanced learning. *Higher Education Research and Development*, **18**:57-75.
- Biggs J 2003. *Teaching for quality learning at University of Buckingham*. London: The Society for Research into Higher Education and Open University Press.
- Black P & Wiliam D 1988. *Inside the black box: Raising standards through classroom assessment*. London: N F E R Nelson Publishing Company.
- Bloom B 1956. *Taxonomy of educational objectives: The classification of educational goals: Handbook I, Cognitive domain*. London: Longman.
- Boughey C 2004. Higher education in South Africa: Context, mission and legislation. In: S Gravett & H Geyser (eds). *Teaching and learning in higher education*. Pretoria: Van Schaik.
- Boulle L 2005. *Mediation: Principles, process, practice*. Sydney: Butterworths.

4 The issue of measurement is clearly implicated in the assessment issue: measurement can be broadly defined as an evaluation of the extent to which a measured behaviour conforms to an external measurement standard; measurement is thus arbitrational in nature. The present article focuses on mediated, non-numerical, formative assessment only (which is strictly separate from measurement per se), and argues that assessment is a mediated practice (thus not arbitrated). Consequently, measurement falls outside the ambit of this article.

5 We acknowledge the need to validate this system with thorough empirical research. However, due to space constraints, this will have to be the topic of a future paper.

- Douglas K & Field R 2006. Looking for answers to mediation's neutrality dilemma in therapeutic jurisprudence. *Elaw Journal*, **13**:177-201.
- Elander J, Harrington J, Norton L, Robinson H & Reddy P 2006. Complex skills and academic writing: A review of evidence about the types of learning required to meet core assessment Criteria. *Assessment and Evaluation in Higher Education*, **31**:71-90.
- Field R 2002. Neutrality and power: Myths and reality. Retrieved from <http://www.mediate.com/articles/fieldR.cfm> on 14 July 2010.
- Francis H & Hallam S 2000. Genre effects on higher education students' text reading for understanding. *Higher Education*, **39**:279-296.
- Glaser R 1963. Instructional technology and the measurement of learning outcomes: Some questions. *American Psychologist*, **18**:519-521.
- Gordon J 2004. The 'wow' factors: The assessment of practical media and creative arts subjects. *Art, Design and Communication in Higher Education*, **3**:61-72.
- Grice P 1975. **Logic and conversation**. In: P Cole and J Morgan (eds). *Syntax and semantics, speech acts*. New York: Academic Press.
- Hammer S 2007. Demonstrating quality outcomes in learning and teaching: Examining 'best practice' in the use of criterion-referenced assessment. *International Journal of Pedagogies and Learning*, **3**:50-58.
- Hussey T & Smith P 2002. The trouble with learning outcomes. *Active Learning in Higher Education*, **3**:220-233.
- Jacobs S 2002. Maintaining neutrality in dispute mediation: Managing disagreement while managing not to disagree. *Journal of Pragmatics*, **34**:1403-1426.
- Killen R 2003. Validity in outcomes-based assessment. *Perspectives in Education*, **21**:1-14.
- Knight P 2001. Complexity and curriculum: A process approach to curriculum-making. *Teaching in Higher Education*, **6**:369-381.
- Knight P & Yorke M 2008. Assessment close up: The limits of exquisite descriptions of achievement. *International Journal of Educational Research*, **47**:175-183.
- Krathwohl D 2002. A revision of Bloom's taxonomy: An overview. *Theory into Practice*, **41**:212-218.
- Maiese M 2005. Neutrality. In: G Burgess & H Burgess (eds). *Beyond intractability*. Boulder, Colorado: Conflict Research Consortium.
- Millman J 1994. Criterion-referenced testing 30 years later: Promise broken, promise kept. *Educational Measurement: Issues and Practice*, **Winter**:19-39.
- Morrow W 2007. *Learning to teach in South Africa*. Cape Town: HSRC Press.
- Niven P 2009. Quit school and become a taxi driver: Reframing first-year students' expectations of assessment in a university environment. *Perspectives in Education*, **27**:278-288.
- O'Donovan B, Price M & Rust C 2000. The student experience of criterion-referenced assessment (through the introduction of a common criteria assessment grid). *Innovations in Education Teaching International*, **38**:74-85.
- O'Donovan B, Price M & Rust C 2004. Know what I mean? Enhancing student understanding of assessment standards and criteria. *Teaching in Higher Education*, **9**:325-335.
- Parker J 2003. Reconceptualizing the curriculum: From commoditization to transformation. *Teaching in Higher Education*, **8**:529-543.
- Perry W 1970. *Forms of intellectual and ethical development in the college years: A scheme*. New York: Holt, Rinehart & Winston.
- Popham W 1993. Educational testing in America: What's right, what's wrong: A criterion-referenced perspective. *Educational Measurement: Issues and Practice*, **Spring**:11-14.
- Price M 2005. Assessment standards: The role of communities of practice and the scholarship of assessment. *Assessment and Evaluation in Higher Education*, **30**:215-230.
- Reed Y, Granville S, Janks H, Makoe P, Stein P, Van Zyl S & Samuel M 2002. [Un]reliable assessment: A case study. *Perspectives in Education*, **21**:15-27.

- Sadler R 2005. Interpretations of criteria-based assessment and grading in higher education. *Assessment and Evaluation in Higher Education*, **20**:175-194.
- Shay S 2004. The assessment of complex performance: A socially-situated interpretive act. *Harvard Educational Review*, **74**:307-329.
- Shay S 2005. The assessment of complex tasks: A double reading. *Studies in Higher Education*, **30**:663-679.
- Shay S 2008. Beyond social constructivist perspectives on assessment: The centering of knowledge. *Teaching in Higher Education*, **13**:595-605.
- Shay S & Jawitz J 2005. Assessment and the quality of educational programmes: What constitutes evidence? *Perspectives in Education*, **23**:103-112.
- Silbey S & Merry S 1986. Mediator settlement strategies. *Law and Policy*, **8**:7-32.
- Singh M 2001. Reinserting the 'public good' into higher education transformation. *Kagisano Higher Education Discussion Series*, **1**:8-18.
- Sizmur S & Sainsbury M 1997. Criterion referencing and the meaning of national curriculum assessment. *British Journal of Educational Studies*, **45**:123-140.
- Smith J 2003. Reconsidering reliability in classroom assessment and grading. *Educational Measurement: Issues and Practice*, **Spring**:26-33.
- Spady W 1994. *Outcome-based education: Critical issues and answers*. Arlington: American Association of School Administrators.
- Strathern M 2000. The tyranny of transparency. *British Educational Research Journal*, **26**:309-321.
- Wiggins G 1990. The case for authentic assessment. *Practical assessment, research & evaluation*, **2**. Retrieved from <http://pareonline.net/getvn.asp?v=2&n=2> on ?? November 2010.