

**Failure of theory: grounding academic standards between rationality and interpretation (0149)**

Academic standards (AS) are at the cornerstone of university education, a fundamental basis for universities' reputation. However, concerns about AS and grade inflation are widespread and, in the UK, have prompted various reports and an investigation by parliament. Yet AS are poorly researched and understood, particularly in their everyday use by academics, managers and those involved in quality assurance. Sadler (2011) notes that the question of what is meant by 'academic standards' is rarely asked, never mind answered and Yorke (2008: 83) describes it as an 'elusive concept'. The nature of standards in assessing complex, higher level, student output is particularly intangible.

Despite this elusiveness, and although statements of standards were rarely articulated in the past, there was an assumption pre-massification, that a 'gold standard' existed; fixed benchmarks which enshrined the standards of the ancient universities (Silver and Williams (1996) 27 & 30). Such a perspective sees standards as independent of the individuals who created or are custodians of them. However, theoretical exposition and empirical evidence is increasingly challenging the robust and reliable nature of such standards as posited by this techno-rational tradition (see for example, Delandshere 2001; Gipps 1999; Shay 2004, 2005; Moss & Shultz 2001).

**Consensus**

There is no doubt that at a fairly general and abstract level, higher education assessors do share a significant measure of agreement about what they are looking for in student work (Warren Piper, 1994). Nevertheless, the hermeneutic critique challenges the potential for shared and stable standards. It argues that, in contrast to the gold standard approach and in the face of evidence that tutors think their standards are the same (Warren Piper 1994, 79), there is growing theoretical argument and empirical evidence for the notion that individuals construct their own 'standards frameworks' (Ashworth *et al.* (2010). These differ depending on academics' values (Shay 2005), on other social worlds that they inhabit (Den Outer, Handley and Price forthcoming, 16), their history (Dobson 2008 A&E 285) and their previous experience (Milanovic, Saville and Shuhong, 106).

On the other hand, and within this epistemological perspective, there is a view that whilst individuals may develop their own standards' frameworks, their 'responses are constituted collectively' (Orr 2010; 15) through activities such as moderation and external examining which help to build an intersubjective consensus. The question is whether this co-constitution of standards is sufficient or is widespread enough to claim a consensus, especially in a diversifying higher education environment? For example, empirical studies (Broad 2003, Bloxham *et al.* 2011, Dobson 2008a, Greatorex 2000, Hawe 2002) have found that assessors use personal standards beyond or different to those stated. In addition, there is a challenge to consensus from the local use of norm referencing (Vaughan, 1991, Shay 2004, Orrell 2008, Orr 2008). Therefore, it can be argued that individualised standards survive regardless of social opportunities to negotiate and construct shared standards and it is not surprising that numerous studies over time indicate low agreement between higher education assessors (Elander and Hardman 2002, Wolf 1995, Bloxham 2009, Leach *et al.* 2001, Shay 2004, Sadler 1987).

The challenge of tacit academic standards has generated a trend towards explicit statements of standards to guide assessors' judgement, particularly in the professions. Statements of standards are also increasingly available to assessors and students in terms of rubrics (marking schemes) and assessment criteria. However studies suggest that such explicit propositions promise more than they can deliver (Hawe 2002, Moss and Shutz, 2001, Sadler 2009a & 2009b). As Broad (2003, 74) states, research shows that 'standards refused to be as solid, stable and portable an entity as

participants wished'. Efforts to make standards explicit such as Subject Benchmarks are not statements of academic standards but 'act as reference points for curriculum design and implementation' (Yorke 2008: 28).

Interestingly, these substantive limitations to the concept of academic standards have hardly entered academic or public debate. Harvey (2002) argues that the external evaluation of Universities has served to legitimise the status quo, focusing more on the process of evaluation rather than the substance of what is being evaluated. Therefore, whilst the stated curriculum and learning outcomes for programmes are now generally in the public domain, one could argue that the judgement of student achievement remains largely unchanged; 'the private preserve of teaching staff' (Coates, 2010, : 10).

The paper will draw on the author's co-research regarding academics learning about standards which found that faculty were torn between techno-rationalist quality assurance imperatives and an implicit grasp of standards as interpretive, co-constructed and tacit. However, the interpretive approach was only embraced to a certain extent with academics continuing to believe in fixed standards and 'right' marks and vesting external examiners with privileged knowledge of AS.

Have we no workable models to help staff bridge this theoretical gulf? On the other hand, perhaps, it is better not to open the Pandora's box of consensus on standards. As Moss and Schultz argue, "No process is ever fully transparent or fully fair and inclusive. Are we better off acting as if we have achieved consensus even if we have not? It is possible that we are." (p64). Conversely, should the academy be seeking a better model of standards for the sake of students and long term reputation. Currently, what confidence can we have that the average academic has the 'assessment literacy' to be aware of the complex influences on their standards and judgement processes? How likely are they to understand the provenance of their own standards and the influence of their background and experience; how strong is the temptation to draw largely on experience as an indicator of what standards should be rather than recognising the potential 'bias' in that approach, or the influence of a particular context or student body or professional experience? Perhaps, it is time to reverse a trend. Safeguarding AS should not rest on creating shared standards through documentation and/or external examining. Instead, perhaps efforts should be directed towards developing models for systematising the social construction of standards through increased intersubjectivity; intra and inter university dialogues.

## References

Alderman, G. (2009) Defining and measuring academic standards: a British Perspective, *Higher Education Management and Policy*, 21 (3), 11-22.

Ashworth, M., Bloxham, S. & Pearce, L. (2010) Examining the tension between academic standards and inclusion for disabled students: the impact on marking of individual academics frameworks for assessment, *Studies in Higher Education*, 35 (2), 209-223.

Bloxham, S. (2009) Marking and moderation in the UK: false assumptions and wasted resources, *Assessment & Evaluation in Higher Education*, 34 (2), 209-220.

Bloxham, S., Boyd, P. & Orr, S. (2011) Mark my words: the role of assessment criteria in UK higher education grading practices, *Studies in Higher Education*, 36 forthcoming.

Broad, B. (2003) *What we really value: beyond Rubrics in Teaching and Assessing Writing* Logan, Utah: Utah State Press

Coates, H. (2010) Defining and monitoring academic standards in Australian higher education, *Higher Education Management and Policy*, 22 (1), 41-58.

Delandshere, G. (2001) Implicit theories, unexamined assumptions and the status quo of educational assessment, *Assessment in Education*, 8 (2), 113-133.

Delandshere, G & Petrosky, A.R (1998), Assessment of Complex performance: Limitations of key measurement assumptions. *Educational Researcher* 27 (2) 14-24.

Den Outer, B., Handley, K and Price, M (unpublished paper, 2010) Standards, situation and self-criticality: exploring situational analysis, grounded theory after the *postmodern turn*, to enhance reflexive practice in higher education research. Oxford Brookes University, UK.

Dobson, S. (2008a): Applying a validity argument model to three examples of the Viva. *Nordisk Pedagogik*, Vol. 28, pp. 332–344 Oslo. ISSN 0901-8050.

Dobson, S. (2008). Theorising the Viva – A Qualitative Approach. *Assessment and Evaluation in Higher Education*, 33(3), 277–288.

Elander, J & Hardman, D (2002: 304) An application of judgment analysis to examination marking in psychology *British Journal of Psychology* (2002), 93, 303–328

Gipps, C. (1999) Socio-cultural aspects of assessment, *Review of Research in Education*, 24, 355-392.

Greatorex, J (2000) ‘Is the glass half full or half empty? What examiners really think of candidates achievement. Paper presented at the British Educational Research Association Conference, Cardiff University, September 7-10 2000

Harvey, L. (2002) Evaluation for what?, *Teaching in Higher Education*, 7 (3), 246-263.

Hawe, E. (2002). Assessment in a pre-service teacher education programme: the rhetoric and the practice of standards-based assessment. *Asia-Pacific Journal of Teacher Education*, 30 (1), 93-106.

Leach, L., Neutze, g & Zepke, N (2001) Assessment and Empowerment: some critical questions. *Assessment & Evaluation in Higher Education* (26) 4, 293-305.

Milanovic, M., Saville, N. & Shuhong, S., (1996) A study of the decision-making behaviour of composition markers in Milanovic, M. & Savill, N (Eds) *Studies in language testing 3: Performance, testing, cognition and assessment*. Cambridge Uni Press, Cambridge, p92-114

Moss, P.A. and Schutz,A (2001) Educational Standards, Assessment, and the Search for Consensus *American Educational Research Journal* 38 (1) 37-70

Orrell, J (2003) Congruence and disjunctions between academics' thinking when assessing and their beliefs about assessment practice, in C. Rust (Ed) *Improving Student Learning: Theory Research and Scholarship*. Oxford: OCSLD

Orrell, J. (2008) Assessment beyond belief: The cognitive process of grading, in A. Havnes & L.McDowell (Eds) *Balancing dilemmas in assessment and learning in contemporary education*. London: Routledge.

Orr, S. (2007) assessment moderation: constructing the marks and constructing the students, *Assessment & Evaluation in Higher Education*, 32 (6), 645-656.

Orr, S. 2008, Real or imagined? The shift from norm referencing to criterion referencing in higher education in *Balancing Dilemmas in Assessment and Learning in Contemporary Education*, eds. A. Havnes & L. McDowell, 1st edn, Routledge, London: 133-144.

Orr, S. (2010) 'We kind of try to merge our own experience with the objectivity of the criteria: the role of connoisseurship and tacit practice in undergraduate fine art assessment. *Art, Design and communication in higher education* 9 (1) 5-19

Sadler, D.R. (2009a) Indeterminacy in the use of preset criteria for assessment and grading, *Assessment & Evaluation in Higher Education*, 34 (2), 159-179.

Sadler, D.R. (2009b) Transforming holistic assessment and grading into a vehicle for complex learning, in G. Joughin, *Assessment, learning and judgement in higher education*. New York: Springer., 45-63.

Sadler, D.R. (1987) Specifying and promulgating achievement standards, *Oxford Review of Education*, 13 (2), 191-209.

Sadler, D.R. (2011) Academic Freedom, achievement standards and professional identity. *Quality in Higher Education*, 17 (1), 85-100

Shay, S.B. (2004) The Assessment of Complex Performance: A Socially Situated Interpretive Act., *Harvard Educational Review*, 74 (3), 307-329; 23.

Shay, S. (2005) The Assessment of Complex Tasks: A Double Reading., *Studies in Higher Education*, 30 (6), 663-679; 17.

Silver, H. & Williams, S. (1996) Academic Standards and the external examiner system, in Brennan, J., et.al. *Changing conceptions of academic standards*. London: Quality Support Centre, Open University., 27-48.

Warren Piper, D (1994) *Are professors professional: The Organisation of University Examinations*, London: Jessica Kingsley Publishers.

Wolf, A. (1995) *Competence-based assessment*. Buckingham: Open University Press.

Vaughan, c (1991) Holistic assessment: what goes on in the raters min, in L. Hamp-Lyons (1991) *assessing second language writing in academic contexts*. Norwood, N.J. Ablex Publishing Corp.

Yorke, M. (2008) *Grading Student Achievement in Higher Education*. London: Routledge.