Accountability in grading: the provenance of standards in a 21st century quality assurance context. (0057)

Bloxham Sue [1], Pete Boyd[1], Susan Orr[2], [1]*University of Cumbria, Lancaster, United Kingdom,* [2]*York St John University, York, United Kingdom*

The massification of higher education has witnessed several decades of increasing regulation and accountability regarding academic standards, for example through the 'Academic Infrastructure' managed by the Quality Assurance Agency in the UK and similar approaches in other higher education systems. In relation to assessment, concerns about standards have expressed themselves in debates about grade inflation, parity of standards across Universities and weaknesses in the systems of degree classification and external examining (UK). The National Student Survey (NSS) in the UK has also heightened institutional anxiety about assessment and feedback. At the heart of these discussions on standards lies the act of grading student work; the professional judgement which embodies our sense of academic standards. Efforts have been made to codify this judgement through artefacts such as institutional or departmental rubrics for marking (grade descriptors) yet the efficacy of these tools to represent the academic 'wisdom' used in grading has yet to be demonstrated. This paper reports on a study of academics' sense of standards as enacted through marking practices. It explores their understanding of academic standards and the roots of their personal frameworks for judgement. It also investigates tutors' sense of accountability for their grading judgements, their use of artefacts (e.g. rubrics) and their attitude to cross tutor and external moderation of their grading including issues of objectivity and subjectivity.

Previous research has examined grading and its limitations (Yorke 2008), the disjunction between lecturer behaviours and pedagogical beliefs in grading (Orrell 2003), grading practices (Grainger et al 2008, Suto & Crisp 2008), moderation effects (Swann & Ecclestone, 1999; Orr 2007), the interplay of subjectivity and objectivity and 'personal interpretative frameworks for judgement' (Shay 2005, Bloxham et al, 2011), learning to mark (Wolf 1995, Jawitz, 2009) the tacit nature of standards and criteria (O'Donovan, Price & Rust, 2008, Sadler 2009a & b) and the interface between lecturers' assessment practices, their values and identity/ies (Orr 2010). This paper builds on this body of work with a particular focus on how tutors position themselves in relation to local artefacts and processes for assuring academic standards in a 21st century context.

A sample of twelve lecturers from two UK universities were asked to 'think aloud' as they graded two written assignments followed by a semi-structured interview. The interview focused on the lecturer's identity, confidence and sense of standards as a marker and their marking practices. The interviewer also recorded field notes. Verbal protocol analysis was selected for its potential to reveal actual, as opposed to espoused, marking practices. Steps were taken to minimize the danger that thinking aloud might influence marking in such a way that students would be disadvantaged.

Both think aloud activity and interview were recorded, transcribed and analysed using a qualitative thematic approach (Ritchie & Lewis, 2003). A coding framework was amended during the process as new themes emerged. Following this, the participants were offered an opportunity to comment on the analysis. The investigation uses a socio-cultural theoretical framework, recognising marking as a socially situated activity (Delandshere, 2001). The data used in this paper is predominantly taken from the interviews.

The theme of accountability emerged in the analysis. Three key elements emerged: who the tutors feel accountable to, how that accountability mediates their grading and feedback activity, and the tension they experience between maintenance of standards and retention of students. The analysis suggests that tutors feel most accountable to their immediate colleagues and external examiners for their grading decisions and these 'peers' are significantly more important than institutional quality assurance demands.   However, the latter comes into play in mediating decisions about the amount and nature of feedback to students, probably in response to NSS pressures. The analysis of tutors' responses about accountability revealed some subtle pressures, for example, to achieve an overall average comparable with other tutors, not to fail too many students and not to be too harsh or too soft a marker.

Tutors were rarely able to express their own sense of standards but the interview data reinforces the view that tutors have, in some way, developed an internalised or embedded set of standards and criteria which they can confidently draw on in their marking. This enabled a high proportion to mark work without immediate reference to text-based artefacts such as assessment criteria or marking rubrics.   Indeed, the strongest finding emerging from the data is a sense that standards are shared and are learnt through collaborative inter-tutor processes. Whilst there is some recognition of subjectivity, for example in the way that criteria are interpreted, there is strong faith in the power of moderation, written criteria and other artefacts to provide adequate consensus. This consensus is particularly strong at the level of local teams and for some it also emerges in relation to their disciplinary community and University. In relation to personal standards, the emphasis remained on whether their marking was 'right'. That is, like many students, they appear to hold the view that there is a correct mark for each piece and moderation helps achieve that 'right' mark if the tutor is unable to.

The analysis supports the view that tutors create or hold their own 'standards frameworks' (Shay, 2005, Bloxham et al 2011) and that these reveal themselves in different ways, including the oral guidance they give through their teaching. For example, a number of tutors highlighted what might be described as 'trigger' qualities in student work which helped them identify the appropriate grade or which side of a boundary work should be graded.

Overall, the study reinforces the view of academic standards as strongly influenced by the local context and informal learning. Whilst staff do not reject the text-based artefacts developed to regularise their grading decisions, the results indicate that they have a very limited role to play in defining or maintaining standards. On this basis, the research reinforces the view that there is a mismatch between existing policy and lecturer practice in relation to the quality assurance of grading. Further research is required to identify more effective ways to secure standards beyond the local context.

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., boyd., P & Orr, S 2011 Mark my words: the role of assessment criteria in UK higher education grading practices . Studies in Higher Education, forthcoming

Delandshere, G. 2001, Implicit theories, unexamined assumptions and the status quo of educational assessment, Assessment in Education, 8, no. 2: 113-133.

Ecclestone, K. 2001, I know a 2:1 when I see it: understanding criteria for degree classifications in franchised university programmes., Journal of Further and Higher Education, 25, no. 3: 301-313.

Grainger, P., Purnell, K. & Zipf, R. 2008, Judging quality through substantive conversations between markers, Assessment & Evaluation in Higher Education, 33, no. 2: 133-142.

Jawitz, J. 2009, Learning in the academic workplace: the harmonization of the collective and individual habitus, Studies in Higher Education, 34, no. 6: 601-614.

O'Donovan, B., Price, M. & Rust, C. 2008, Developing Student Understanding of Assessment standards: a nested hierarchy of approaches, Teaching in Higher Education, 13, no. 2: 205-217.

Orr, s. (2010) 'Being an artist you kind of, I mean, you get used to excellence': Identity, values and fine art assessment practices. International Journal of Art and Design Education forthcoming

Orr, S. 2007, assessment moderation: constructing the marks and constructing the students, Assessment & Evaluation in Higher Education, 32, no. 6: 645-656.

Orrell, J. 2003, Congruence and disjunctions between academics' thinking when assessing and their beliefs about assessment practice, Improving Student Learning: Theory Research and Scholarship, ed. C. Rust, OCSLD, Oxford: 186.

Ritchie, J. & Lewis, J. (eds) 2003, Qualitative Research Practice: a guide for social science students and researchers., Sage, London.

Sadler, D.R. 2009a, Indeterminancy in the use of preset criteria for assessment and grading, Assessment & Evaluation in Higher Education, 34, no. 2: 159-179.

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

Shay, S. 2005, The Assessment of Complex Tasks: A Double Reading., 30, no. 6: 663-679; pp. 17.

Suto, W.M.I. & Crisp, V.'.G., J. 2008, Investigating the judgemental marking process: an overview of our recent research, Research matters, 5: 6-9.

Swann, J. & Ecclestone, K. (1999) 'Litigation and learning: tensions in improving university lecturers assessment practice', Assessment in Education, 6 (3), pp.357-375.

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

Yorke, M. 2008, Grading Student Achievement in Higher Education, Routledge, London.