UNIVERSITY OF
GLOUCESTERSHIRE

This is a peer-reviewed, final published version of the following document:

Calvert, Ben (2005) Demystifying Marking: reflections on developing and using grade descriptors: Case Study. Learning and Teaching in Higher Education (1). pp. 93-97.

EPrint URI: https://eprints.glos.ac.uk/id/eprint/3612

Disclaimer

The University of Gloucestershire has obtained warranties from all depositors as to their title in the material deposited and as to their right to deposit such material.

The University of Gloucestershire makes no representation or warranties of commercial utility, title, or fitness for a particular purpose or any other warranty, express or implied in respect of any material deposited.

The University of Gloucestershire makes no representation that the use of the materials will not infringe any patent, copyright, trademark or other property or proprietary rights.

The University of Gloucestershire accepts no liability for any infringement of intellectual property rights in any material deposited but will remove such material from public view pending investigation in the event of an allegation of any such infringement.

PLEASE SCROLL DOWN FOR TEXT.

Demystifying Marking: reflections on developing and using grade descriptors

Ben Calvert
University of Gloucestershire, UK

Marking is a necessarily contentious activity. It is not always possible to arrive at bull's-eye conclusions about the merit of an individual piece of work, particularly in the arts, social sciences and humanities where studies of culture, society and history do not yield principles or general rules that can be tested for accuracy. If it were possible, double marking would be unnecessary and regulations allowing students to appeal their marks would become more or less redundant. But I use the words 'necessarily contentious' because marking should compel tutors to engage with one another about this difficulty, to keep on asking what it is that they are looking for in student work. Although we may not agree on the finer detail, we should welcome common ground where we can find it. It is vital to double-check that we know which skills and knowledge we are testing, that these are set at an appropriate level, that we all know when they have been met, and that they cohere within a curriculum. However, in modular programmes it is challenging to reach shared understandings when working lives can be isolated and disconnected.

In 2003, tutors in the Field of Media Communications at the University of Gloucestershire, attended an away day to improve marking. The session combined reflection on existing practice with planning for improvements. A key aim was to place marking in context. It was hoped that staff would better understand the subject's response to nationally disseminated standards in UK universities (QAA, 2002) and to locally-derived position statements on them (i.e. the Programme Specification for our course). The significance of these processes would become apparent by making them relevant to 'coalface' delivery, in this case assessment and marking. Throughout the day the idea that setting and grading assessments should be an expression of wider curriculum aims, as outlined in our Programme Specification, was reinforced. However, provisional discussions revealed that recently-recruited staff were largely unaware of national and institutional initiatives and saw them as rather distant and mysterious. Related developments, such as the Quality Assurance Agency's (QAA) guidelines for assessment (QAA, 2000) were also unfamiliar. A principal function of the session was for staff to familiarize

themselves with them by reading a document that tracked connections between benchmarking, programme specifications, level descriptors, learning outcomes, assessment criteria and grade descriptors.

The away day was organized as a mock QAA inspection meeting. Questions ranged from issues about student progression (e.g. 'What would you expect a model first year, second year and third year student to demonstrate in terms of skills and knowledge?'), to concerns about specific theoretical content (e.g. 'How do students' skills of textual analysis progress across three years?'). Staff were also asked how they differentiated between skills and knowledge within various levels (e.g. 'What would an A grade student in year three demonstrate in textual analysis compared to a B, C, D or fail student?'). This revealed information on student attainment that is often kept 'in the heads' of staff as implicit knowledge. Varied understandings also led staff to recognize the need for a degree of consensus in order for the curriculum to be coherent.

The team then provided examples of how our expectations are articulated by writing learning outcomes for a mode of textual analysis at all three levels, demonstrating progression. They then derived marking criteria related to these, informing students of the elements against which they would be assessed. Finally, grade descriptors were composed detailing the skills and knowledge required in order to score grades in each category. An example of practice in another subject was disseminated to assist in this. Previously, bespoke grade descriptors had been absent from our modules in respect to specific items of assessment. This is not unusual, with Rust (2002, p.147) arguing that although the use of learning outcomes is common, 'the linkage between the outcomes and the coursework essay, exam or whatever is tenuous at best, and almost always implicit'. For us, such descriptors existed only in generic form in the course guide, with statements applying across all three levels. The team agreed that these were inadequate for mapping attainment levels in individual assessments testing a variety of skills.

The team composed descriptors according to a grid system, with sets of statements applying to each marking criteria. An example of format is given in Figure 1, relating to an examination on a compulsory level two module. It is worth noting that individual written comments on student performance are also included on the definitive descriptor sheet.

At the end of the session a holistic understanding of the relationship between benchmarks and our curriculum, flowing down to individual assignments, outcomes, criteria and grade descriptors, had been achieved. It was agreed that from September 2003 module guides would include statements linking the module to wider programme aims, along with bespoke grade descriptors for all assessments. An agreed template for these guides secured consistency of practice. The

Figure 1: Example of Grade Descriptors

Marking Criteria: The exam answers will be assessed against the following criteria. All are given equal weighting:

i. Clarity of and specificity of the response to the question with evidence of a sustained and balanced argument

ii. Ability to evaluate and critique theoretical concepts and perspectives effectively and persuasively

iii. Ability to arrive at an informed and persuasive conclusions

iv. Clarity and fluency of expression

Grade Descriptors: marks will be awarded according to the following grid:

Marking Criteria	70+	60-70	50-60	40-50	30-40	Below 30	Mark
i	Response fully addresses the question and draws on a range of advanced sources to construct a balanced and sustained argument	Response addresses the question and draws on relevant sources to construct a balanced argument	Response addresses the question using relevant sources with evidence of a basic ability to construct an argument	Response addresses the question but may lack relevancy in places, may not fully utilize relevant academic sources and lacks a sustained argument	Response may address the question indirectly and arguments tend to be intuitive or descriptive with few relevant academic sources consulted	Response to the question partial or unclear with unconvincing, poorly sourced arguments	55
ii	Theoretical positions fully and clearly explained, demonstrating an ability to critique persuasively on the basis of evidence. May contain evidence of original thinking or synthesis	Theoretical positions clearly explained, demonstrating ability to critique persuasively using evidence	Theoretical positions explained and critiqued but explanations may not be full or may have some inaccuracies, or critique lacks detail or persuasiveness	Theory explained at a basic level, but little evidence of critique. May contain some theoretical inaccuracies	Partial or inaccurate exposition of theory with only an intuitive ability to critique demonstrated	Little or no engagement with relevant theory	62
iii	Theoretical debates and evidence carefully drawn together to construct a focused argument that convinces the reader fully of the writers understanding. May contain evidence of original thinking	Theoretical debates and evidence summarized effectively. Evidence of an ability to construct a persuasive argument	Theoretical debates summarized clearly, with a basic argument or position taken	Theoretical debates summarized but with little evidence of an informed argument or a convincing summary position being taken	Theoretical debates inadequately summarized. Any arguments likely to be intuitive or descriptive	Little or no evidence of a theoretical conclusion. Entirely intuitive or descriptive	64
iv	Impressive command of language, including spelling and grammar. A fluency of style and an ability to make complex debates clear is in evidence. Work well structured with clear bridges and links built between key points	Clear command of language, including spelling and grammar. Style is generally fluent and an ability to make points clearly is in evidence. Work is soundly structured	Good command of language, including spelling and grammar, although style may lack fluency. Soundly structured, but may lack some clear bridges and links between key points	Sound command of language, including spelling and grammar. May contain some basic mistakes or lack fluency or a clear structure in places	Basic command on language but several mistakes in spelling and grammar are evident, and work lacks fluency or a clear structure	Poor command of language with many basic mistakes of spelling and grammar and little fluency or structure	58

template was derived to ensure that curriculum aims and the learning and assessment strategy of modules spoke to each other in a process of 'constructive alignment' (Biggs, 2003). It was also hoped that descriptors would be useful for students in formative assessment by directing attention to the meeting of module outcomes. For staff they would draw attention to common issues within assessments and focus minds on whether or not agreed outcomes had been met. In both cases, the aim was to demystify marking.

Currently, our use of this approach is being monitored. Preliminary views have revealed a balance of positive feedback and genuine concerns. Whilst all staff felt that descriptors focused their attention more clearly when marking, and in particular made double marking more transparent, some felt that this was at the expense of an overall evaluation of a particular piece of work and its merit. Some would read the work and allocate an 'instinctive' mark and then apply the descriptors to fit this impression. They believed that certain assignments 'felt like' a 65% or 50%, for example, but argued that this mark might not emerge via strict adherence to the grade descriptors. As a result, they ended up marking in much the same way as before and adapted the descriptors to their existing marking style. The view that marking (and perhaps teaching and learning more widely), is a complex activity that cannot be adequately reproduced in a tighter regime of marking was also expressed. In an appraisal of the use of learning outcomes, Hussey & Smith (2002) have argued that, when misused, they perform a managerial function as an aspect of the commodification of learning. This reflects less their use for deepening learning and more 'the desire to audit and monitor the performance of those involved'. It is possible that some staff may see the use of grade descriptors in the same way. Whether or not this was the case within the team, the exercise of judgment when marking had been confronted and in some cases changed.

This led to further discussions about the way in which the criteria were written. Was it better for them to be tight or loose in order for some discretion to be allowed? The provision of an additional catch-all descriptor, based on a general impression of the work was also debated. This could allow staff to weight the work according to both the tightly written descriptors and their impressionistic sense of its merit. Others felt that this would remystify the process again and would run against the spirit of using the descriptors in the first place. Student responses have also been mixed. In module evaluations some have welcomed the use of the 'shaded grid' method as it focuses clearly on what they need to do to score particular grades. However, they do not see tutors shading the grid as 'feedback', which they perceive only as a personal written statement. Although the team do write individual feedback as well, some students felt that there was an issue of balance between generic responses (albeit useful ones) and those that are more personalized.

At the moment decisions about whether to modify their use are on hold as it is felt that time is needed to adjust to a new mode of marking. Staff recognized that they have the power to determine the phrasing of the descriptors as long as they reflect the appropriate marking criteria and learning outcomes. If staff wished to write descriptors that encouraged perhaps more elusive factors (e.g. originality or a creative approach to a piece of work) then this was fine, as long as students were aware that this was an expectation and as long as marking criteria and module outcomes required those elements to be assessed. At the end of the meeting it was emphasized that descriptors do not necessarily diminish the professional judgement of staff, they just require them to exercise it differently, in a way that is more transparent and meaningful for students. In summary, the team is continuing to use grade descriptors and is generally persuaded by their value, contingent upon modifications being discussed at the end of a full academic year.

References

Biggs, J. (2003) *Teaching for Quality Learning at University*, 2nd ed, Buckingham: Society for Research into Higher Education and Open University Press.

Hussey, T. & Smith, P. (2002) The trouble with learning outcomes, *Active Learning in Higher Education*, vol. 3, no. 3, pp. 220-233.

Qaa (Quality Assurance Agency for Higher Education) (2000) *Code of Practice – Quality and Standards in Higher Education. Section 6 'Assessment'* [online]. Available from http://www.qaa.ac.uk/public/Cop/COPaosfinal/contents.htm.

Qaa (Quality Assurance Agency for Higher Education) (2002) *Benchmark Statement for Communication, Media, Film and Cultural Studies* [online]. Available from http://www.qaa.ac.uk/crntwork/benchmark/phase2/communications.htm.

Rust, C. (2002) The impact of assessment on student learning, *Active Learning in Higher Education*, vol. 3, no. 2, pp. 145-158.

Keywords

Grade descriptors, marking, benchmarking, constructive alignment

Biography

Ben Calvert's background is in the social sciences and he has recently published work on television studies. He has undertaken a funded project on effective dissertation supervision and is currently undertaking research on developing a research culture in first year undergraduate students.