

## **Second-Marking and the Academic Interpretative Community: ensuring reliability, consistency and objectivity between markers**

James Bennett  
Department of Applied Social Sciences  
London Metropolitan University

**Keywords:** *inter-marker reliability, interpretative community, media studies, Hourly-paid-lecturers (HPLs)*

### **Introduction**

A central issue to any attempt to improve and monitor inter-marker reliability revolves around the need to comply with the *QAA Code of Practice's* stipulation that 'institutions must have transparent and fair mechanisms for marking and moderating marks' (Section 7). Various research theories and practices have been designed by different academic practitioners to ensure this goal is met, but these often focused on different moments in the assessment process. Thus White's (2000) article on monitoring argues for an attention to the end of the marking process by making second markers write a report focusing on the behaviour of the first marker, rather than on the student's individual paper as in double-marking. In contrast, Sandhu's (2004) article 'The usage of statement banks as a means of assessing large groups' suggests that not only can objective criteria be made known through assessment schemes published at the start of the assessment process, but also through the use of standardised feedback statements that clearly indicate where a student's piece of work lies within the assessment guidelines.

For the newly appointed academic, arguably the response to this plethora of options and theories is to recognise that no one method alone can offer solutions to the problems of fairness, transparency and reliability consistently encountered between markers. Although I feel this conclusion, is valid something about all these techniques felt rather unsatisfying; namely, that all approaches still relied on markers interpreting one document or another in the same manner. That is, markers had to all belong to the same 'interpretative community' not only in order to agree on marks, but also to experience the kind of 'systematic disagreement' that White identifies as an area of healthy debate in academic assessment and easily resolved by process monitoring (2000). I am largely in favour of White's support of monitoring but want to consider here some of the pre-conditions necessary for a system of inter-marker reliability to work consistently. In this article, I therefore examine the centrality of inter-marker reliability to good academic practice. I then go on to outline debates about the interpretative community, sketching these theories on to

the issues involved in university assessments, particularly the challenges of integrating hourly paid lecturers (HPLs) into this community.

### **Inter-marker reliability**

The importance of developing an interpretative community for assessment is underscored by the QAA's requirements for fair and objective assessment schemas set out above and, more particularly by London Metropolitan University's *Assessment Framework* stipulation that methods for marking students' work:

*allow everyone involved in the process including students ... and the markers themselves, to have confidence that the marks awarded are a fair reflection of the quality of ... the work submitted (2004: 19)*

However, creating a fair assessment process that is transparent and open to all of those involved is not a simply a process of publishing assessment guidelines and marking schemas. Of equal, if not greater importance, is the need to ensure that such criteria, along with essays themselves, are interpreted in broadly similar ways.

This is a problem that Hand and Clewes sought to examine by conducting research to 'understand how tutors went about deciding the grades that they allocated to a student's dissertation work' (2000: 7). As they suggest, there are broadly two approaches to ensuring inter-marker reliability and fairness: the criteria approach and a more 'holistic approach'. Whilst the former claims to avoid marking discrepancies between markers by setting out clearly worded criteria that are 'spelled out and agreed among staff, and communicated to students', the latter perceives written essays (in this case a dissertation) as best evaluated as a 'qualitative piece of work' whereby any attempt 'to reduce the work to specific elements, each with their own criteria, would diminish the judgments of tutors' (2000: 7). However, as their study goes on to attest, ensuring consistency and reliability is never such a clear cut choice between these two methods. Crucially, their study provided 'no understanding of exactly *how* those articulated values are put into practice ... values that we look for within dissertation work are, by their very nature, qualitative' (*ibid*: 16). Their work suggested that although assessment criteria are designed to make the assessment and marking process a more transparent process, certain terms remain vague and open to subjective interpretation. These are discussed by Webster et al who suggest terms like 'discuss', 'logical structure', or requiring students to think 'critically', 'conceptually' or 'laterally' are all interpreted differently by individual markers and students alike.

Faced with such problems, Hand and Clewes (2000: 18) are forced to conclude that 'perhaps the practice of double marking existed to test out the judgements used and to prevent differences in interpretation?'. Their uncertainty of relying on the process of double-marking to solve these problems is evinced by the caveat they add at the end of the article. Having considered the problems of double-marking (for example

those identified by Ecclestone, 2001, such as bullying, deference to senior colleagues etc), the authors go on to suggest that double marking will nevertheless be valuable where 'the process includes an open exchange of views about values, standards and qualities ... [when] faculty may begin to move towards common understandings'.

A selection of the academic work that takes into account how different readers of a text might come to such common understandings is discussed below followed by an examination of own practice.

### **The interpretative community**

Stanley Fish has provided some of the founding legal theory approaches to how judges and lawyers should interpret legislation and previous judicial decisions. During the 1980s, Fish engaged in public debate with Owen Fiss regarding the ability of judges to be objective in their interpretation and application of legislation. Whilst Fiss argued that texts constrain the readers to such a large extent that there is no room for personal opinion or bias, Fish suggested that the meaning of any text could only be deciphered within the bounds of an 'interpretative community'. Fish argues further that the consistency of judges' understandings of legislation has been largely a result of their socialisation and training into a professional body, however he also suggests that the constraints of this community are not absolute. Rather, a shared interpretative community does not guarantee a new objectivity but instead is still subject to individual interpretations and discovery of values within the text (Andrews, 1998).

Fish's work echoes Roland Barthes' work on cultural theory. In Barthes' later post-structuralist writings in 'The Death of the Author' (1968[1977]) and *S/Z* (1974[1990]), he argued that neither the author of a text, nor the words on the page themselves were the source of meaning of any text. Instead, he suggested that the text's meaning only came alive when the reader invested these dead words of the text with meaning. As Barthes famously proclaimed, 'the birth of the reader must be at the cost of the death of the Author' (1968[1977]). For example, in marking a media studies essay the individual marker will bring their experience of marking other essays, as well as the appropriate media studies reading material, to the process of marking this particular essay. This might include experience of marking essays at different institutions where the calibre of students is substantially higher or lower than at London Metropolitan. As a result, the marker might think this essay is particularly good or poor, based on the intertextual reading they bring to the script. Of course, the argument advanced by writers such as Webster, Hand & Clewe and Sandhu (*op.cit.*) is that marking criteria prevent the marker from bringing a completely subjective assessment to the essay. However, as Barthes' work suggests, these criteria themselves are open to interpretation and the work of the reader.

Such a position is, of course, untenable for a fair, objective, transparent and objective assessment system. However, the Birmingham School of Cultural Studies provides another viewpoint which might be productive ground for thinking about how shared understandings can, and are, reached. Hall (1980) suggests that texts are not simply polysemic, but rather their multiple meanings are limited by patterning constraints, such as class, gender, age etc. More specifically, in the case of ensuring inter-marker reliability, these patterned constraint might be thought of as the development of professional practices that form an interpretative community, which I want to close my discussion with here by relating these ideas to my experience in teaching modules in media studies at London Metropolitan University.

### **The media studies marking interpretative community**

So what is our interpretative community for marking and what are the structured limitations on our ability to read and understand both assessment criteria and marking schedules as well as an individual students' essay. I work in a faculty where staff come from a variety of different backgrounds in terms of race, class, gender and most importantly academic discipline. However, we must form an interpretative community in order to set assessment items that each other can understand and with which we can fairly monitor the marking of colleagues. There are a number of ways in which this is done, which I only have space to name here: a collegial environment; the promotion of shared research, such as through research groups; the use of student feedback; the frequency and quality of discussions about particular modules and the degree course itself. I want to concentrate on this final point for a moment here, as it is of interest to my final concern: the integration of HPLs into the interpretative community.

Because the teaching pool for media studies is relatively small and many of us take research leave to write articles, books and develop media applications, there is a necessity for each of us to teach modules that have been written and/or developed by others in the department. Of course, this causes some problems that are discussed below but importantly leads to an 'exchange moment', which if carefully managed can result in open discussion about what has/hasn't worked in terms of the module's syllabus, assessment schemes and criteria. Such a system ensures a general awareness of the degree syllabi and the standard of work expected and received across a wide variety of modules. As a result, rather than a total disagreement about how an individual marker has graded papers, before the module begins, colleagues are able to talk about what kinds of response to assessment schemes have been received in the past and, should disagreement arise, refer back to these as a way of resolving disagreements.

However, with a small pool of staff of the kind described above, there is a need to delegate a variety of teaching sessions to hourly-paid lecturers (HPLs), a situation to be found in many 'new universities'. Such lecturing staff are less-likely to be part of

the interpretative community and their contacts with the University may consist only of the area's academic leader and the person who previously convened the module they are now teaching. Because such HPLs are not usually strongly affiliated with the University they are less likely to interpret assessment schemes and student essays in the same way as others in the department and this is a major challenge to which there are no simple solutions. Clearly for this system to be effective they need to be involved in the processes and forums outlined above but not only is this difficult to facilitate because HPL appointments are often made very close to course start dates, it is also often resisted by HPLs themselves for a variety of personal, professional and social reasons. The result is that not only are HPLs unlikely to challenge potential employers when it comes to monitoring or second marking, they are also more open to challenge themselves as they are usually less aware of what constitutes appropriate 'benchmarks' for marking/feedback etc.

## **Conclusion**

A stronger interpretative community might reduce the chance of deferral, bullying and, in turn, obviate some of the problems identified with White's ideals of monitoring. Monitoring, if done properly and without fear of bullying etc, can provide an opportunity for all markers to become more aware of their decision making processes and this can only be to the good for all stakeholders in the process. However HPLs need to be included in this process, which must occur at the start of their appointment rather than at the moment when grades are being debated. That problem aside, monitoring and the existence of a strong interpretative community can serve to reduce marker subjectivity and provide an important patterning constraint on the interpretation of individual essays and assessment criteria.

In summary, Ecclestone (2001) suggests that the phrase 'I know a 2:1 when I see it' can be interpreted as a 'genuine espoused theory', yet ... there is a fine balance between the genuine ability to recognise quality of work ... and erratic interpretation'. By allowing the marker to become more self-aware through monitoring and also more constrained - not only by texts but also through the 'socialisation into a professional body' of shared 'training, practice and value' that leads to 'internalised ways of reading and understanding' texts - monitoring regimes and the development of 'interpretive communities' may go some way to supporting the academic's divine ability to spot the all important 2:1.

## **References**

Andrews, N. (1998). 'Wormes in the entrayles: the corporate citizen in law?', Murdoch University Electronic Journal of Law, vol. 5(2):  
<ftp://cleo.murdoch.edu.au/pub/elaw/issues/v5n2/andrews521.txt>

- Barthes, R. (1968[1977]). 'The Death of the Author' in Image-Music-Text. London: Fontana Press.
- Barthes, R. (1990[1974]). *S/Z*. (Trans. Richard Miller). Oxford: Blackwell Publishing.
- 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*, Vol. 25(3); pp. 301-313.
- Fish, S. (1980). *Is There a Text in This Class? The Authority of Interpretive Communities* Boston, Mass.: Harvard University Press.
- Fish, S. (1982). 'Working on the Chain Gang: Interpretation in Law and Literature' *Texas Law Review*, vol. 60: 551
- Fiss, O. (1982). 'Objectivity and Interpretation', *Stanford Law Review*, vol. 34: pp.739.
- Hall, S. (1980). 'Encoding/Decoding'. Hall, S. Hobson, D. Lowe, A. & Willis, P. (Eds). *Culture, media, language*. London: Hutchinson. pp. 128 – 139.
- Hand, L and Clewes, D (2000) Marking the difference: an investigation of the criteria used for assessing undergraduate dissertations in a business school. *Assessment & Evaluation in Higher Education*, 25 (1) 5-21.
- Read, B. & Francis, B. (2003). 'University lecturers' construction of undergraduate writing: a gender analysis'. *Investigations: In University Teaching and Learning*, vol. 1(1); pp. 59-62
- Sandhu, S. (2004). 'The usage of statement banks as a means of assessing large groups'. *Investigations: In University Teaching and Learning*, vol. 2(1); pp. 49-55.
- Webster, F, Pepper, D and Jenkins, A (2000) Assessing the undergraduate dissertation *Assessment & Evaluation in Higher Education* 25 (1); 71-80.
- White, R. (2000). 'Double marking versus monitoring of examinations'.  
[www.prs-ltn.leeds.ac.uk/philosophy/discussions/white.html](http://www.prs-ltn.leeds.ac.uk/philosophy/discussions/white.html)

## **Biographical note**

James Bennett is Senior Lecturer in Media Studies at London Metropolitan University and is currently completing his PhD on Digital television, interactivity and public service broadcasting. He has previously been the recipient of an international scholarship to study at the University of Madison-Wisconsin.

Email: [j.bennett@londonmet.ac.uk](mailto:j.bennett@londonmet.ac.uk)