

# Assessing quality of feedback in online marking databases: An opportunity for academic professional development or just Big Brother?

Shane McKenzie

Department of Criminology  
University of Melbourne

This paper explores possibilities for utilising student assignment feedback, written by academics and stored in an online marking and results system called OMAR at the University of Melbourne, for quality assurance of feedback and to provide professional development for academics. The paper draws upon the results of a survey of staff and students using OMAR in 2003, which suggested that the quality of feedback is related to the individual academic's commitment to the task more so than any assistive technology (McKenzie, 2003b), and while this is intuitive it is yet to be tested. The mere thought of assessing the quality of feedback given to students via data and text mining of online marking databases has already taken some academics beyond their traditional comfort zones for performance appraisal. The possibilities raise legal and ethical issues in relation to privacy, consent, accountability and appropriate uses of the data that might be generated. On the other hand, taking advantage of analysing such databases could present a better means of providing feedback to academics on one of their most common tasks, which would benefit both academics and students.

**Keywords:** feedback, online marking, assessment, professional development, knowledge discovery and data mining, text mining

---

## Introduction

*The pile of essays in front of you seems endless as you complete yet another feedback sheet, award a grade and pick up the next essay. Scanning quickly over the pages, you realise that is not very good, and sighing, settle in to read it in detail. We have all had that sinking feeling during marking student work. How it is approached from there can differ dramatically. Some have been tempted to write derisive or sarcastic comments in response to student work that does not meet expectations. Others may be tempted to simply award the grade and move on without insightful comment to the next assignment, which might be more engaging to the examiner. A rarer breed of academic is able to discern the positive aspects of this mediocre piece of work and phrase comments that identify the problem areas, provide constructive feedback and maintain the student's desire to learn.*

New academics, whether they are first time lecturers or tutors, may never have written assessment feedback before. Learning to write useful and appropriate feedback for students has traditionally been an apprenticeship process, with staff learning from more experienced teachers in their discipline. This can be a hit and miss process; as training aimed specifically at workshopping written feedback is in general unavailable. Increasingly, quality assurance of assessment, including feedback, is important for universities. Which begs the question, "How do we know if the feedback provided by academics to students is effective and helps them learn?"

This paper explores some possibilities of utilising student assignment feedback, written by academics and stored in an online marking and results system, called OMAR and developed at the University of Melbourne, for quality assurance of feedback and to provide feedback and professional development to academics.

## Quality and effective feedback

What constitutes effective or quality feedback for students? Many studies have looked at how teachers do or should write feedback to students (See eg, Sommers, 1982; Straub and Lunsford, 1995; Christiansen, 2004). Advice provided to academics at the University of Melbourne by the Centre for the Study of Higher Education suggests that feedback needs to be timely, informative, and be focused more on

development of the student than judgment of them, by making suggestions for improvement (James, McInnis and Devlin, 2002). Assessment is not the end of the learning process. They argue that “Really conscientious marking involves pointing out each individual flaw in logic or inadequacy of treatment” (James, McInnis and Devlin, 2002: 3) and providing feedback best suited to assist that individual student, which is sensitive of their feelings (James, 1994; Beattie and James, 1996). It is also important for academics to avoid discriminatory language in their feedback to students, which might stereotype, label negatively, sensationalise, trivialise or denigrate the student (Equal Opportunity Unit, University of Melbourne, 2002).

This can sound like a tall order for academics, in light of increasing administrative burdens and demands on their time. Especially, when it has been suggested recently that students do not read or assimilate all the feedback provided by academics on their work (Graham Gibbs cited in Mills, 2004). Kennedy and Judd’s (2004) analysis of audit trails in multimedia software also suggests that students do not use feedback as one would expect. This all suggests there is a fine balance to be found between too much, too little or poorly focused feedback, and it is worthwhile not just to the student, but for the academic that they get this balance right.

### **OMAR and feedback for students**

OMAR (the Online Marking and Results system; <http://www.omar.unimelb.edu.au>) was designed and programmed by the author in the Department of Criminology at the University of Melbourne to assist with his own and others’ marking. It is web based software that permits academics to create and customise online templates for performing their marking of student assignments and returning that feedback to students. OMAR also facilitates electronic submission and some other administrative tasks. OMAR is different to other “online assessment” software, as it does not provide functionality for online testing, but focuses instead upon the marking and provision of feedback for traditional student assignments, whether they are essays, oral presentations, tutorial attendance, exams or something else.

Since the first prototype was developed in late 2000, OMAR has been used by eighty subjects in the Arts and Vet Science faculties, by 179 staff and just over 4,500 students in 111 semester cohorts. The latest, daily usage statistics are available from <http://www.omar.unimelb.edu.au/docs/usage.html>. At the time of writing this paper, OMAR stored just over 12,000 assessments for 239 separate tasks assigned to students.

The ideology behind my development of OMAR was that “*machines can do the work, so that people have time to think*” (B(if)Tek, 2000). In this sense, “work” meant stapling feedback sheets to assignments, calculating grade distributions, alphabetising stacks of essays and returning the assignments to students, and “thinking” concerned providing effective feedback to students. If the routine administration could be streamlined it would provide more time to consider the students’ work, and this has been the author’s experience of the system thus far.

OMAR provides many different ways academics can provide feedback to students in their marking templates. From defining criteria and then assigning either marks or using Likert scales, writing comments, and returning files to the students, whether these are additional resources for the whole class or a student’s electronic submission, which can be marked up using other software, such as a word processor. Further details can be found in the online User Guide, which is available publicly at <http://www.omar.unimelb.edu.au/docs>.

OMAR was designed so that examiners can view the feedback written by their colleagues earlier than in the traditional marking process. Traditionally, examiners meet up after they have completed all their marking to do cross marking, when adjusting comments or grades means twice the work and so not all student assessments are cross marked. OMAR promotes earlier cross marking and consistency as examiners can access the assessments written by others teaching on their subject as soon as they are written into the system, and can generate on the fly statistics about their grade distribution. This also provides an avenue for new academics to learn from their peers about how to write appropriate and useful feedback.

## Feedback from staff and student users of OMAR

The author carried out an online survey of staff and student users of OMAR in late 2003, asking them to comment upon their experiences of the system, including questions about the amount and quality of the feedback compared to their subjects that did not use OMAR (McKenzie, 2003b). The results from this survey prompted the idea of analysing the feedback stored in OMAR, in particular the comment from a student that “*Quality is a property of the marker, not of the system.*”

This would seem to be commonsense, indicating that an academic’s commitment to the task of providing feedback is more important than any characteristic of an online marking system. However, is it an accurate observation? Does using an online marking system make no difference to the quality of feedback provided to students? Examiners need to fill in a marking template, which might provide and enforce a structure not used previously in their marking. This by itself might promote improved consistency in the type and quality of feedback provided to students. Or, there is also the potential that online marking could sever the direct link between comments and student work if the system is not used sensibly, making feedback less useful to students (McKenzie, 2003a).

In the 2003 survey, some students indicated they had received more feedback in the assessments they received through OMAR:

Excellent - same or more feedback than usual

... the amount of feedback OMAR have provided is adequate (marking with fair, poor, etc) so that we know where the assignment's lacking off, and we can still have comments from our tutor. This is really good since the manual feedback often not in such great detail

I got more feedback in my OMAR subject than in my other subjects.

However, others complained they had received less feedback:

... feedback is poor compared with comments written on the essay. much more difficult to relate the comments to the essay, particularly wrt grammer or punctuation or even general organisation. overall comments using omar tend to be briefer and less helpful. it surely must be easier to mark the essay itself rather than take the additional step of entering assessment into omar.

Again, on the amount of feedback provided, it was noted that:

Unfortunately it is dependant on the staff who wish to really use the opportunity to give complete feedback that the system does work. Some do not make use of returning the essay with full comments which I find would be a great advantage to student who wish to improve. Perhaps it could be suggested to staff to always utilise this facility and not just the comments section that appears with the marks page.

For staff, their experiences of online marking were positive overall:

I like being able to read other's comments. It gives me ideas for how to provide tactful feedback. It also gives me an idea of how I am travelling in comparison with other markers. (Tutor)

A great means of cross checking for consistency (Full time academic)

The benefits to staff extended to easier identification of students with learning difficulties:

Greatly improved identifying students with learning problems by making other markers opinions accessible to the group as a whole. (Full time academic)

Given the sizeable database of student assessments stored in OMAR, there is an opportunity to test this empirically in comparison to feedback sheets not written using OMAR. What are the characteristics of the feedback provided using OMAR? Do staff using OMAR provide more or less feedback of better or worse quality? Are their comments more or less developmental or judgmental? Does the feedback make links to the course material and learning objectives? Or are the comments generic? The following sections discuss the possibilities of undertaking such empirical analysis of the OMAR database to answer these questions.

## Technical considerations

The OMAR database design stores identifying information, such as names and student numbers, separately to other elements of the marking templates. Because of this, it would be quite easy technically to extract for analysis elements, such as comments written to a student, anonymous of both the student's and examiner's identifying information. The characteristics of the marking templates created in OMAR could also be extracted anonymous of their particular subjects or academics.

Ideally, in the long term, analysis of feedback to students would be automated, given the large numbers of students and assignments per subject. The aim is not to create more work for academics. One possibility would be to use knowledge discovery and data mining techniques upon the database to analyse marking template characteristics (Brankovic and Estivill-Castro, 1999). The classification process in data mining works on the principle of inputting a training set, that is a set of example cases and their classes, and outputting a classifier that will assign classes to new cases (Brankovic and Estivill-Castro, 1999). Once data has been classified it can be clustered and have predictive modelling performed on it. How could a training set be constructed for the OMAR database? They suggest judgements would either need to be made about existing feedback or on dummy assessments. A panel of senior academics or educational experts could make these judgments.

The analysis of the feedback written to students in comment items on the templates could perhaps be analysed using text mining. The Wikipedia Encyclopaedia defines text mining, also known as intelligent text analysis, text data mining and knowledge discovery in text, as:

... the process of extracting interesting and nontrivial information and knowledge from unstructured text. Text mining is a young interdisciplinary field, which draws on information retrieval, data mining, machine learning, statistics, and computational linguistics. As most information (over 80 percent) is stored as text, text mining is believed to have high commercial potential value.

Some options available now include using SAS Text miner or Predictive Text Analytics from SPSS. The promise is that text mining can deal with unstructured data and will generate knowledge the academic did not think to ask for initially. However, these tools rely upon a skilled expert analysis of the findings, due to the ambiguous nature of language and discipline specific information that could be contained in the texts (Robb, 2004). For the purposes suggested in this paper, the text mining process would need to be incorporated as background functionality of the OMAR software.

Another possibility for achieving more timely, systematic feedback for the academic would be to facilitate student rating of the feedback they receive for an assignment through OMAR using a poll, as is found increasingly at the end of commercial software support web documents, which ask, "How useful was this document to answering your question?" Students could be asked, "How useful was this feedback to improving your learning?" or similar. Consideration might need to be given to diverting the student from their actual grade to focus on the issue of quality feedback.

Individualised feedback about their online marking could then be provided through the OMAR interface to the academic, describing what they do and directing them to advice on best practice assessment techniques if needed, such as teaching and learning resources cited above and advice on constructing better marking templates.

## Potential benefits, legal and ethical concerns

Is the analysis of databases of student assignment feedback a legitimate avenue for quality assurance and professional development or simply another example of Big Brother style surveillance by employers? Each academic to whom this author suggested the idea of analysing the OMAR database to provide feedback to them indicated that was "a scary concept", which obviously placed them beyond their traditional comfort zones. Perhaps this is a positive indication of the need to test the quality of feedback?

Universities need to be accountable for their processes of assessment. Individual academics who coordinate subjects bear the everyday responsibility for the assessment of their students, regardless of whether they delegate marking to tutors. Quality assurance of feedback to students needs to occur more

regularly than just for new staff, as suggested by the opening scenario to this paper. The potential benefits of analysing the OMAR database includes feedback to academics allowing them to seek professional development or a sense of satisfaction knowing their feedback is of quality, improved student learning and hence student satisfaction with their courses, which could translate into increased enrolments and revenue for universities.

Fox (2001: 11) argues that the public regards the current level of surveillance and dataveillance as essentially benign due to its fragmented, decentralised and distributed nature between the public and private sectors. The fear of a Big Brother behind it all seems unfounded. However, workplace monitoring of performance is a hot issue in the Australian private sector, and this paper is exploring a form of it for academics. Resistance to over surveillance is framed generally in terms of privacy issues (Fox, 2001), and the counter argument is based upon legitimacy. Brankovic and Estivill-Castro (1999) identified several threats from mining databases for knowledge discovery that should be considered, such as stereotyping, generation of misinformation, and breach of privacy through disclosure or inappropriate combination of results with other results. They suggest that researchers undertaking data mining analyses should have need to know access to individual data (Brankovic and Estivill-Castro, 1999: 94).

Universities might be entitled to use and analyse assessments written by staff depending upon their intellectual property regulations. To be ethical, it is suggested each staff member give that informed consent before analysis takes place. Further, under the *Information Privacy Act 2000* (Vic) and the *National Privacy Principles*, information such as student grades and feedback would be considered personal information, although not sensitive information for the purposes of the Acts. The analysis of such information by a university for the purpose of quality assurance and professional development would most likely constitute a related secondary purpose, and so be permissible.

Analysing the feedback simply in terms of frequency of positive or negative comments is insufficient, as the need to point out faults with the students' work will be dependent upon the quality of that work. One might be tempted to address this issue as whether the quality of an academic's comments reflect the quality of the student work. From the author's experience of reading his own and others' feedback sheets, there is sometimes a tendency to provide more feedback for students who have failed than for students who have performed well. In one sense, Christiansen (2004) identified this as academics covering themselves against comebacks by disgruntled students, however it could also point to difficulties in providing suggestions for improvement to high achievers. High achieving students often need feedback as much as those who have performed less well, and there is also the need to justify higher grades.

Whether and how the results of such analyses should be used in relation to performance appraisal of academics is controversial. At my university, quality of teaching surveys of students are not supposed to be used for performance appraisal due to the subjective nature of the anonymous feedback. Would analysis of academic feedback to students by text mining be more objective? Software designers would need to be accountable for any software processes used to make these analyses, similar to those tests used increasingly to judge the validity of computer forensic software and digital evidence by the courts (Casey, 2001). Would we, as users of this technology, know how the information was generated and be able to verify its accuracy?

Quality of teaching results are often communicated to students, but only in aggregate, statistical form. Should the results of analysing academic feedback be communicated to students? Again, probably only in aggregate form to protect the privacy of students, if they were to be disclosed at all. It would be unethical to classify extracts from the database and then quote these to other staff and students as examples of either good or poor practice, where this would identify the author or the student involved.

## Conclusions

Graham Gibbs (1999) suggests that there are parallels between the processes of improving research and improving teaching. Peer review should be an important part of improving the assessment process, including the peer review of feedback to students. "What is judged by peers is valued and what is valued is usually pursued with vigour and intelligence" (Gibbs, 1999). OMAR already provides a mechanism for improved peer review of feedback written for students. Technology, such as text mining, could provide opportunities for formalising and improving this process, however its implementation and use needs to be carefully monitored, just as is the research process. Prerequisite to performing analyses of an academic's

feedback to students is that policies are agreed as to the appropriate use of data generated, that ethical and legal concerns are managed and that appropriate feedback is provided to the academic, with training where necessary.

## Acknowledgements

The author wishes to express his thanks to the Department of Criminology and the ITMM Committee of the Faculty of Arts for supporting the development of OMAR and the writing of this paper. Thanks to # anonymous referees for comments on an earlier version of this paper.

## References

- Beattie, K. and James, R. (1996). *Assessing Essays*. Centre for the Study of Higher Education, University of Melbourne. [verified 21 Oct 2004] [http://www.cshe.unimelb.edu.au/downloads/revise\\_handouts/Assessing\\_rev2.doc](http://www.cshe.unimelb.edu.au/downloads/revise_handouts/Assessing_rev2.doc)
- B(if)tek (2000). *Machines work, 2020*, CD recording. Melbourne: Murmur records. MATTC105
- Brankovic, L. and Estivill-Castro, V. (1999). Privacy issues in knowledge discovery and data mining. In C. Simpson (Ed), *Flow on effects of Information Technology on quality of life and the environment*. Proceedings of the first Australian Institute of Computer Ethics conference. [verified 21 Oct 2004] <http://crpit.com/confpapers/CRPITV1Wahlstrom.pdf>
- Casey, E. (2001). *Handbook of Computer Crime Investigation*, New York: Elsevier.
- Christiansen, R. (2004). Critical discourse analysis and academic literacies: My encounters with student writing. *The Writing Instructor*. [verified 21 Oct 2004] <http://www.writinginstructor.com/essays/christiansen-all.html>
- Fox, R.G., (2001). Someone to watch over us: Back to the Panopticon?, *Criminal Justice*, 1(3), 251-277. [verified 21 Oct 2004] <http://crj.sagepub.com/cgi/framedreprint/1/3/251>
- Gibbs, G. (1999). Improving teaching, learning and assessment, *Journal of Geography in Higher Education*, 23(2) July, 147-155.
- James, R. (1994). *Assessment*. [http://www.cshe.unimelb.edu.au/downloads/assessment\\_rev2.pdf](http://www.cshe.unimelb.edu.au/downloads/assessment_rev2.pdf)
- James, R., McInnis, C. and Devlin, M. (2002). *Assessing Learning in Australian Universities: Ideas strategies and resources for quality in student assessment*. Centre for the Study of Higher Education, University of Melbourne. [verified 21 Oct 2004] <http://www.cshe.unimelb.edu.au/assessinglearning/>
- Kennedy, G.E. & Judd, T.S. (2004). Making sense of audit trail data. *Australasian Journal of Educational Technology*, 20(1), 18-32. <http://www.ascilite.org.au/ajet/ajet20/kennedy.html>
- McKenzie, S. (2003). OMAR: Improving the Online Marking Process. Invited paper and poster presented to the *Multimedia & Educational Technologies for Teaching and Learning Enhancement (METTLE)* conference, University of Melbourne, Nov 5.
- McKenzie, S. (2003). *Results from the Feedback survey of Staff and Student Users of OMAR 2003*. [verified 21 Oct 2004] <http://www.omar.unimelb.edu.au/docs/surveyresults.html>
- Mills, R. (2004). Learner support: Developments in open and distance education and their implications for traditional educational institutions. Paper presented to *The Open and Distance Learning Association of Australia's Professional Development Seminar series*, University of Melbourne, 2 July.
- Robb, D. (2004). Taming Text. *Computerworld*, June 21, 40-41.
- Sommers, N. (1982). Responding to student writing. *College Composition and Communication*, 33, 148-56.
- Straub, R. and Lunsford, R.F. (1995). *Twelve Readers Reading: Responding to College Student Writing*. Cresskill, NJ: Hampton Press.
- Wikipedia Encyclopaedia (nd). Text Mining. [verified 21 Oct 2004] [http://en.wikipedia.org/wiki/Text\\_mining](http://en.wikipedia.org/wiki/Text_mining)

**Mr Shane McKenzie**, Lecturer, Department of Criminology, The University of Melbourne VIC 3010  
Email: [shaneem@unimelb.edu.au](mailto:shaneem@unimelb.edu.au) Web: [www.crim.unimelb.edu.au/staff/shaneem.html](http://www.crim.unimelb.edu.au/staff/shaneem.html)

**Please cite as:** McKenzie, S. (2004). Assessing quality of feedback in online marking databases: An opportunity for academic professional development or just Big Brother?. In R. Atkinson, C. McBeath, D. Jonas-Dwyer & R. Phillips (Eds), *Beyond the comfort zone: Proceedings of the 21st ASCILITE Conference* (pp. 623-628). Perth, 5-8 December.  
<http://www.ascilite.org.au/conferences/perth04/procs/mckenzie.html>

Copyright © 2004 Shane McKenzie

The author assigns to ASCILITE and educational non-profit institutions a non-exclusive licence to use this document for personal use and in courses of instruction provided that the article is used in full and this copyright statement is reproduced. The author also grants a non-exclusive licence to ASCILITE to publish this document on the ASCILITE web site (including any mirror or archival sites that may be developed) and in printed form within the ASCILITE 2004 Conference Proceedings. Any other usage is prohibited without the express permission of the author.