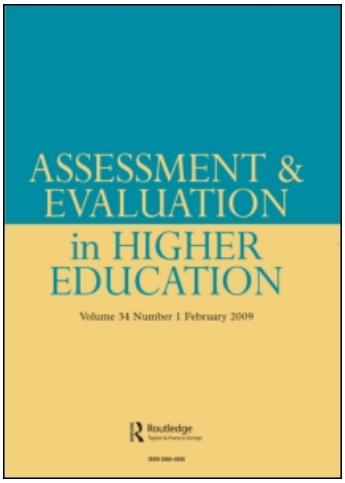

Assessment & Evaluation in Higher Education
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713402663

From copying to learning: using exemplars to engage students with assessment criteria and feedback

Karen Handley[a]; Lindsay Williams[a]
[a] Oxford Brookes University, Oxford, OX33 1HX, UK

First published on: 02 October 2009

To cite this Article Handley, Karen and Williams, Lindsay(2009) 'From copying to learning: using exemplars to engage students with assessment criteria and feedback', Assessment & Evaluation in Higher Education,, First published on: 02 October 2009 (iFirst)
To link to this Article: DOI: 10.1080/02602930903201669
URL: http://dx.doi.org/10.1080/02602930903201669

Assessment & Evaluation in Higher Education
2009, 1–14, iFirst Article

From copying to learning: using exemplars to engage students with assessment criteria and feedback

Karen Handley* and Lindsay Williams

Oxford Brookes University, Oxford, OX33 1HX, UK

Feedback is central to pedagogic theory, and if feedback is to be effective, students need to engage with it and apply it at some point in the future. However, student dissatisfaction with feedback – as evidenced in the National Student Survey – suggests that there are problems which limit student engagement with feedback, such as their perception that much of their feedback is irrelevant to future assignments. This article reports on a study which sought to enhance engagement by giving students exemplar assignments annotated with feedback *before* submission of their final assignments. This was done by providing an online facility where students could view exemplars and post comments or questions to tutors and peers on a discussion board. The exemplar facility was highly valued by students, although there were no quantitative effects such as an increase in students' assignment marks when compared with the previous cohort. The article reflects on possible reasons for this result and discusses ways to improve the exemplar facility, for example by facilitating dialogue between tutors and students. The article concludes with lessons learned about how to construct exemplars, and considers how exemplars might also be used within marking teams to improve consistency of marking.

Keywords: feedback; exemplars; assessment criteria; e-learning

Introduction

Student engagement is acknowledged to be 'a highly desirable goal with positive outcomes for all parties' (Bryson and Hand 2007, 354). The process of engagement is subtle, however, and can be difficult to operationalise for the purposes of empirical research. Engagement is more than the sum of observed behaviours inside or outside the classroom, with peers or tutors; it also involves '[mental] interaction with content' (Moore 1989). Furthermore, cognitive engagement may occur at a surface or deeper level (Entwistle 1988) – yet may manifest itself in the same way to observers. For example, a student who is reading tutor feedback on an assessed essay may be scanning it quickly for a justification of the mark; or may be relating the feedback comments to the assignment criteria and then inferring principles of good practice to be applied in the next assignment. In the later case, the 'deeper' engagement with feedback has at least the potential to lead to changes in behaviour or understanding.

We propose in this article that if our interactions with students are to be pedagogically effective, students must engage with them. This is particularly the case with formative assessment feedback which relies for its effectiveness on being applied at

*Corresponding author. Email: khandley@brookes.ac.uk

ISSN 0260-2938 print/ISSN 1469-297X online

DOI: 10.1080/02602930903201669
http://www.informaworld.com

some point in the future. However, a key issue for many students – and perhaps the main reason behind their lack of engagement with feedback – is that they perceive much of their feedback to be irrelevant to future assignments and modules.

Student perception of the 'irrelevance' of feedback may arise for a number of reasons. Firstly, students may be unable to understand the feedback. This hypothesis is supported by studies of student comprehension of *written* feedback, which is probably the most common medium for delivering feedback in the social sciences. Chanock (2000), for example, found considerable variability in the ways students interpreted the words 'description' and 'analysis', partly because of different traditions across disciplines, and partly because students (especially first-year undergraduates) 'are not insiders to the disciplines they study' (97).

Another, more systemic, problem is that students may decide *not* to engage with feedback if they see no link from one module/tutor/topic to the next. This behaviour may occur if students see modules as unique rather than progressive, and where they perceive no value in applying feedback to subsequent assignments. Lack of engagement with feedback is a serious problem because unless students engage with it, feedback cannot be effective, which means that time spent by tutors preparing it is wasted (Gibbs and Simpson 2004).

A third issue is that unless students understand the assessment criteria for their work, that is, the context for the feedback – they cannot fully interpret and 'decode' (Nicol 2008) the feedback. This limits feedback effectiveness as well as student engagement with it. Feedback is rarely intended as a new piece of content – as something which is comprehensible and usable in its own right. Instead, feedback is conventionally understood as a mechanism for helping students see and then close a performance 'gap', and develop their self-assessment skills (Sadler 1989; Boud 1995). This gap is the difference between the student's own work and an idealised performance indicated by the assessment criteria and standards, for example in the form of a criteria grid. However, if students cannot grasp the meaning of assessment criteria/standards, feedback intended to guide them on achieving those standards may be incomprehensible and therefore ineffectual. Using a sporting analogy, can high jumpers really understand why they have 'failed' if they cannot see the high-jump barrier? and can the coach's verbal feedback ever compensate for the high-jumper's inability to see the target?

These issues are closely related to one another: unless students anticipate that feedback will help them understand their 'performance gap', engagement will be poor. The inter-relationship between engagement and perceived effectiveness is enduring and complex. In the research presented in this article, we investigated both issues whilst focusing on problems of student disengagement which were seen as the most pressing, and which influenced our research design and use of exemplars. However, as we analysed the research findings, we identified that a deeper problem was students' difficulty in developing their tacit understanding of how to interpret assessment criteria, which was vital in order to fully interpret their feedback. This aspect is elaborated in our discussion section.

The problem of student engagement with feedback

Lack of student engagement with feedback (and their perception of its effectiveness) has no single, simple solution. This is disappointing, but should not be surprising given that learning from feedback is usually an iterative, ongoing process and not just

a one-off event. Nevertheless, some insights may be gained from considering the problem from a temporal perspective: for example, how does the timing of feedback influence student engagement? In particular, how do we shift the timing of feedback so that students appreciate its usefulness?

There are several options for 'time-shifting' feedback, for example:

- feedback-on-drafts (where feedback is given on students' draft assignments to enable improvements to be made to the final submission) (e.g. Price et al. 2007); and
- exemplars (where feedback is annotated on exemplars of previous assignments, or where students generate feedback by comparing exemplars with the assignment criteria) (e.g. Orsmond, Merry, and Reiling 2000, 2002).

The first approach involves time-shifting feedback so that it is given mid-module, *before* submission of their forthcoming assignment. This could be achieved by designing an assessment structure where formative feedback is given on *draft* assignments (Millar and Szwelnik 2007), allowing students to reflect on the feedback before developing the final version. The drawback of this approach is that – in the absence of modifications such as a policy of only summatively marking the final assignments – it will generate higher marking workloads for tutors.

Another way to shift the timing of feedback, which is the subject of this article, is to engage students mid-module with feedback written for previously marked exemplar assignments completed by students in earlier cohorts. In this way, students are positioned as vicarious learners (Bandura 1986). By giving students the exemplar assignments with associated marking criteria and feedback, students can be helped to interpret the feedback in the context of the exemplars, and *then* translate their learning into the context of their current assignment. Another potential benefit of this approach is that students may take a more active role in their learning than might be the case if they are merely 'told' the criteria requirements of the assignment.

The aim of the research reported here was to investigate the utility of exemplar assignments as a means of conveying tacit knowledge about criteria standards to students in a large module taught across 14 seminar groups. The research is reported in the following way. We begin by discussing some of the prior research on the pedagogic use of exemplars. Then, in order to provide a context for the study, we give background information on the module and students, and the nature of the exemplar facility. This is followed by a description of the research design and methods used to elicit student behaviours and attitudes. We then present the research findings, and finally, we discuss our interpretation of the findings and some implications for future pedagogic practice. The latter section includes a discussion of the use of exemplars within tutor-marking teams.

Research on exemplars

Exemplars have been described as 'key examples chosen so as to be typical of designated levels of quality or competence. The exemplars are not standards themselves but are indicative of them … they specify standards implicitly' (Sadler 1987, 200). Used in this way, exemplars can convey to students the tutor's tacit knowledge about what criteria actually mean (Polanyi 1973). Exemplars may also be helpful in modules with large teaching teams where marking consistency depends on shared understandings of

explicitly worded assessment criteria (Rust, Price, and O'Donovan 2003; O'Donovan, Price, and Rust 2004; Price 2005). From the perspective of students, exemplars are engaging for instrumental as well as developmental reasons: *instrumental* in that exemplars may promote student insights about how to improve their own assignments; and *developmental* in that exemplars help students refine their understanding of their discipline and how to communicate within it.

Although exemplars are not *models*, students may see them in this way. The term *model* reminds us of the idea of 'model answers' used as targets which students should aim for. Models often work through observation and imitation – and here there may be problems: firstly, that imitation is not the same as learning; and secondly, that imitation may lead to plagiarism. Nevertheless, exemplars potentially share these problems *if used* as models without intervening dialogue to help students draw out principles and insights about assessment criteria (e.g. see Huxham 2007).

Although research on exemplars is a relatively small field, a number of important studies exist which informed the research reported in this article. In two seminal studies, Orsmond, Merry, and Reiling (2000, 2002) used exemplars during modules which required students to create scientific posters for their biology assessments. The aim of the initial study (Orsmond, Merry, and Reiling 2000) was to improve students' understanding of assessment criteria through tutor/student discussions leading to joint development of marking criteria. However, the authors found that this activity did *not* necessarily improve tutor/student agreement regarding the *standards* relevant to those criteria. They suggested that one reason for students' difficulties was that they were unfamiliar with what a scientific poster 'looked like' and were therefore unable to fully contextualise their tutor's comments about marking criteria (Orsmond, Merry, and Reiling 2002, 310). The authors proposed that providing exemplars of other students' work might overcome this difficulty (Orsmond, Merry, and Reiling 2002, 310).

The second exemplar study (Orsmond, Merry, and Reiling 2002) demonstrated the utility of using exemplars in a module involving student-constructed marking criteria, followed by student self- and peer-assessment of completed science posters. Exemplars of previously created posters illustrating the topic of histology were shown to students to focus the discussion of criteria-construction. The posters were presented simply as illustrations of different styles without tutor feedback about the merits of individual posters, and without divulging the grade awarded to each piece of work. Instead, the exemplars prompted discussions about the different *qualities* of the posters, which were translated into a set of criteria to be used in self- and peer-assessment of the students' own work. An important finding from the study was that the use of exemplars in conjunction with student-constructed marking criteria reduced the amount of 'difference' between student and tutor marks. However, in the post-course evaluation, students 'seemed to be less enthusiastic' about the process of criteria-construction, perhaps because of the demanding nature of such a task (Orsmond, Merry, and Reiling 2002, 320). It seems that students find the process of learning from exemplars 'difficult'.

The Orsmond study used a relatively small group of students (n = 22), and was able to engage students in comprehensive discussions about the exemplars and assessment criteria. The extent to which the approach is transferable to other modules depends on a number of factors including the discipline of study and the level and number of students. For programme directors with large modules, the scalability of

this approach is particularly relevant. The ability to use exemplars in a large module across multiple seminar groups was a key question framing the present study, which is discussed in the remainder of this article.

Empirical context

Research aims

This study examined how exemplars could be used in a large module involving multiple seminar groups (14 each semester, with 20–28 students in each group), run by a multi-disciplinary team of 10 tutors and lead by two co-module leaders (the authors of this article). In particular, the authors wanted to establish whether students could learn from exemplars posted onto the university's virtual learning environment (VLE) (WebCT), where online discussion about the criteria, assignments and feedback was to be moderated by module leaders. A key difference compared with the Orsmond design was that the exemplars were annotated with feedback. However, unlike the Orsmond study, students did not construct marking criteria, and there was no in-class discussion of exemplars.

The aims of the study were to explore the following questions:

- How do students interact with a databank of marked assignments (including browsing the exemplars, and interacting with the VLE discussion forum)?
- To what extent do students come to understand the assessment requirements?

The module and students

The module selected to investigate the exemplar facility was an undergraduate, second-year module called Methods of Enquiry. The module helps students develop their information sourcing and evaluation skills, learn the fundamentals of research design, and practise data collection methods such as questionnaires and interviews. Methods of Enquiry is a large module (400 students in Semester 1; 325 students in Semester 2). Forty per cent of the module marks relate to a coursework assignment of 1700 words with two components: a report on a topical business problem; and an evaluation of information sources selected for the assignment. The assignment is completed by student pairs and submitted in week 5 of the semester.

Exemplar facility

Exemplars were collected by asking students from the previous cohort to contribute their de-identified assignments. Two assignments were offered and available for Semester 1, and an additional two assignments were offered for Semester 2. The exemplars were similar in structure to the assignment which students were about to complete, but had a different topic focus. The exemplars were de-identified for reasons of confidentiality, annotated with feedback, transformed into PDF documents, and uploaded into the WebCT contents page.

For one of the contributed assignments available in Semester 2, an interactive page was created in WebCT using CourseGenie software. This enabled us to present parts of an exemplar assignment followed by self-test questions prompting students to reflect on the quality of the text. Some questions were text-free (with a pre-written

comment from staff as the 'answer') and others were constructed as multiple choice questions.

Students were advised in week 2 that exemplars were available to help with preparation for their coursework on the topic of organic food. Students were also reminded that they could post any queries on the linked WebCT discussion thread. All queries (e.g. asking for clarification about the feedback given on an assignment) were to be answered by the module leaders, with to the aim of deepening students' understanding of the marking criteria, and the qualities tutors look for when marking assignments.

Research design

The research questions presented earlier were investigated using data from an anonymous student questionnaire, informal discussions with students and usage statistics available from WebCT.

The questionnaire was hosted at www.surveymonkey.com, a user-friendly service providing online surveys. To encourage responses, only four questions were asked: (1) Overall, how useful was the exemplar facility (not at all; not useful; fairly useful; very useful); (2) What did you like? (3) What didn't you like? (4) How could the facility be improved. Responses to Question 1 provided a numerical evaluation of the exemplar facility, Questions 2–4 provided qualitative responses.

Statistical data were collected by tracking the number of 'hits' per exemplar assignment, using WebCT's page tracking functionality. Data were collected over a number of weeks preceding the hand-in date in Week 5 of the module. In addition, we collected anecdotal data from conversations with students.

Analysis

Our analysis of the survey and page-tracking data showed that students made significant use of the exemplar facility, and that they found it very useful. Total hits in Semester 2 numbered 1361, that is, just over four hits per student (with a cohort of 325). A similar pattern was evident for Semester 1. In both semesters, the greatest use of the facility occurred in the week preceding the hand-in deadline.

The survey response rate was 15% in Semester 1 (63 out of 400 students). This is a relatively low response rate but is not surprising: students generally respond poorly to surveys because of other commitments and survey fatigue. Nevertheless, the large size of the cohort means that the absolute number of responses is sufficient to warrant further analysis. It is also important to acknowledge potential non-respondent bias. This bias often manifests as excessively positive or negative responses (e.g. where only dissatisfied customers bother to record their attitudes in customer questionnaires). In the student questionnaire, however, the range of positive and negative open-text responses suggests that a broad spectrum of perspectives were recorded, and gives some reassurance that the natural variability in students' opinions is reflected in the 63 responses. We have no reason to believe that non-respondent bias is a material problem for the interpretation of data. Technical problems with the survey facility in Semester 2 reduced the response rate and limits the usefulness of that data.

An analysis of the Semester 1 data shows that 73% of respondents found the exemplar facility 'very useful'; 24% found it 'fairly useful'; and 3% responded that it was 'not useful' or 'not at all useful'. The response was thus overwhelmingly positive. An

analysis of student responses for Question 2 ('what did you like?) show that almost half cited the benefits of seeing the structure and layout of the exemplar assignments. For example:

> I could see the layout and structure so I had an idea about how to put it together. (S36)
>
> It gave a first impression of what the assignment should look like. (S46)
>
> I finally understand the structure and format of the assignment. (S13)

Other students mentioned the reassurance gained from seeing what they should be aiming for. One students also mentioned the motivating effect of seeing a good assignment:

> [I liked] being able understand exactly what is expected. (S1)
>
> A reference point. (S16)
>
> I was inspired by the superior quality of the assignment. This was highly motivating. [I wanted to] match or beat the quality and standard. (S17)
>
> We could check our own work against the layout to make sure we were on the right lines. (S29)

A quarter of students referred to the feedback, for example:

> It was useful to see what comments were made about the assignment, pushing us into the correct direction. (S5)
>
> I liked the useful hints from lecturers about how to write and how to improve the work. (S14)
>
> Good to see what the module leader was looking for. (S55)

In response to Questions 3 and 4 (what didn't you like? and what could be improved?), students made fewer comments, or answered with statements such as 'I liked everything'. However, a few suggestions were made. The main one, mentioned by about one-tenth of respondents, was that it would be useful to include examples of poor assignments. One student, recognising the potential for plagiarism, said: 'I need to remember not to copy it.' Another student made the following suggestion:

> The examples were good. Perhaps if they are handed out in class without comments and then discussed as a group – the strengths and weaknesses – before receiving comments from the marker, it would greatly improve our idea of what is expected. (S3)

Disappointingly, no students used the discussion facility in WebCT to ask questions or post comments about the exemplar assignments or feedback. A comparison of coursework marks is shown in Table 1. It was disappointing to note that student marks did not significantly improve following introduction of the exemplar facility. Indeed, the lack of improvement in marks was unexpected given students' positive comments about how the exemplars helped in producing their coursework. There are several possible explanations for the lack of apparent impact on marks.

Table 1. Comparison of coursework marks across three cohorts.

	Number of students	Average coursework mark (out of 40)	Standard deviation of coursework mark
Semester 2 (2005–2006)	354	23.5	5.4
Assignments database introduced for 2006–2007			
Semester 1 (2006–2007)	405	22.0	5
Semester 2 (2006–2007)	311	22.3	5.6

Firstly, there is always the possibility when conducting quasi-experiments in a natural setting that the results are 'contaminated' by the influence of intervening variables. This possibility is endemic in pedagogic research which acknowledges the multiplicity of interacting variables on student learning and performance, and the difficulties of isolating their influences on individual performance (Salomon 1992). In this particular case, changes in module leadership from 2005–2006 to 2006–2007 lead to several adaptations and developments in content, emphasis and teaching approach including (but not limited to) the addition of the exemplar facility. These changes may have generated a cohort-level impact on student marks which masked the beneficial impact of the exemplar facility (or vice versa).

Secondly, our reflections on the operation of the exemplar facility, supported by comments from students and the teaching team, suggest that our implementation strategy was deficient in several ways. There is scope for substantial improvements which we discuss next. In the spirit of learning through continual development, we hope that our own experiences will provide insights which others can build upon.

Whilst acknowledging that average marks did not improve, we can nevertheless point to other data showing a positive through relatively weak correlation between students' 'hits' on WebCT, and their coursework marks. In this context, 'hits' include all WebCT resources including – but not limited to – the exemplar assignments. We do not have data at a granularity which permits us to match exemplar hits with student marks, but we suggest that data on total hits are an adequate proxy for exemplar hits since the latter accounted for 57% of all hits. Specifically, there is a statistically significant relationship between a student's coursework mark and his or her total WebCT hits, $r = 0.28$, p (two-tailed) < 0.01.

Discussion

The positive response from students, both in terms of their high usage of the example facility, and their positive ratings in the questionnaire, suggest that they found it a very valuable resource. Students commented that they liked being able to see the structure and layout of a previously marked assignment, and to read the feedback in order to understand what was required, and how the assignment could have been improved. Nevertheless, our analysis raises a number of questions.

The high volume of comments in favour of seeing 'structure and layout' raise an issue about whether the benefits of the exemplar facility could be achieved simply by improving the assignment brief. Student comments suggest that they did not quite understand what was expected of them, and wanted to see templates and examples to encourage them 'along the right lines'. However, whilst the magnitude of students' comments on this topic were unexpected, it is perhaps not surprising that second-year

undergraduate students are still struggling with the basic structural elements of report-writing, and need examples *in addition* to a description of the assignment brief. However much tutors 'explain' what is required, this is not necessarily sufficient: students need to 'see' how that explanation translates into a real example.

A concern raised in the mind of the module leaders, based on post-semester discussions with some students, is that students accept the exemplar feedback *at face value* and do not necessarily understand it in the way intended when written. This seems to be especially problematic because students were not the authors of the original assignments, and so the feedback was not targeted to them in the first place. A related issue is that students may infer incorrectly that unless there is a feedback comment suggesting improvements, an assignment is 'perfect'. This is of course a dilemma faced by all markers, but it may be worsened if students believe that a marked assigned uploaded into an exemplar facility is a 'perfectly-and-fully-marked assignment', with comments on ALL problem areas. In practice, markers may only identify salient areas.

Some students, when asked about their own coursework, appeared to have misinterpreted the feedback, or made erroneous assumptions about what was good or bad about the exemplar. One student, for example, complained about his low mark and came to one of the authors for clarification. On reading his coursework, it became evident that he had made statements which were incorrect, for example stating that 'all sources were peer reviewed', when in fact he had used articles from non-peer-reviewed magazines and websites. When this statement was queried, it was clear that the student had picked up from an exemplar that 'peer review' was an indicator of good quality sources, but he had failed to pick up that peer review only relates to certain types of journals and a few other sources. He responded by saying 'but it said on the example that peer review was a good thing …'.

The student's comment implies that he was merely copying what he thought had attracted good marks (and in this case, did so incorrectly). His action raises the question of whether 'good' exemplars are taken as *models* by students, and whether their availability facilitates plagiarism. On the other hand, if students need to see models of structure and layout in order to develop their academic literacy and develop report-writing skills, do the benefits outweigh the dangers of copying?

The potential problem of misinterpretation was compounded by student resistance to using the discussion forum. We had expected students to read the exemplars and the feedback, reflect on what they had read, and post queries onto the forum allowing module leaders to clarify misunderstandings. This step in the learning process did not happen. This may have been for a number of reasons, for example shyness, not wanting to appear ignorant, or not knowing how to frame an 'appropriate' question. To some extent, student behaviour can be interpreted using a theoretical framework of situation learning theory (Lave and Wenger 1991). Students were accustomed to pedagogic methods which involved lectures or class activities during workshops and seminars. They were also accustomed to reading material 'given' by the tutor and which students would assume to be good or 'models'. Lecture attendance and reading are fairly passive activities in comparison with workshops and seminars. By posting the exemplars on WebCT, students seemed to interpret them as 'reading material' to be absorbed without much reflection. This behaviour was not what had been hoped for by the module leaders, who expected a more discursive response where students would comment in the discussion forum about the feedback – what it meant, how the feedback could be acted upon, and so on. However, this invitation to commit to a written debate was a step too far for the students. It seems

that no one wanted to disagree with the given feedback, and no one wanted to display their ignorance by asking questions publicly in an online forum. Judging by the response from student S3, however, some students would have welcomed a traditional discussion in class led by the tutor. This proposition is supported by research conducted as part of an FDTL[I] project on ways to engage students with assessment feedback (Price et al. 2007). During FDTL focus groups with students on what good formative feedback 'looks like', students commented on the advantages of seeing exemplars:

> It really helps to see past questions, see a past person's work, see what was a first and what was a third. [But we want a conversation] – not just to download them.

Our discussion has so far focused on student engagement. The findings suggest that students engaged with the exemplars and the feedback annotations which they considered useful to guide the writing of their assignments. In one sense, the exemplar facility was successful: by time-shifting the feedback so that it was used to annotated exemplars, students were engaged, and most questionnaire respondents described the facility as 'very useful'. However, our informal conversations with students revealed the extent of their difficulty in understanding the explicitly written assessment criteria. This lack of understanding seemed to limit students' ability to learn from the non-discursive exemplar facility, and may explain why average marks did not rise following its introduction. What was needed a dialogic process by which tutors could share their tacit ways of interpreting explicitly written criteria, so that students could begin to see those criteria embedded in the exemplars.

Two further questions are raised by our analysis: What makes for a 'good' exemplar and how can exemplars be used in large teaching teams to facilitate a sharing of tacit knowledge among a group of markers?

What makes a good exemplar?

As indicated earlier, students often find it difficult to abstract from exemplars what the different criteria 'look like', especially in the absence of formative dialogue with peers or tutors. Quantity is not necessarily the answer. Providing more and more exemplars works only if we can depend on students' powers of inductive reasoning to elicit tacitly what tutors cannot say explicitly; namely what the criteria and levels/standards definitively mean. However, we know from the previous research that inductive reasoning is a difficult process and is fraught with mental traps, such as the tendency to remember only the most recent exemplars (Gick and Holyoak 1983; Detterman and Sternberg 1993).

Nevertheless, students can engage with what is *visible* to them (Sadler 1987). What, then, can we *make* visible? This question presumes some degree of tutor planning in the selection of exemplars (or even in their re-construction, which we discuss later). Qualities of an assignment which we may want to 'make visible', and about which we can give formative feedback, may involve *criteria* (such as 'appropriate use of Case Law to demonstrate the principle of precedence'), or *standards* (such as 'discussion appropriate at masters' level'). Qualities which are visible to tutors may be invisible to students. A key skill for academics is the ability to stand in the shoes of the student audience: What problems or illustrations of good academic practice will students recognise, or can be made visible through dialogue?

If students are presented with poor assignments, they may be unable to isolate and then analyse the qualities they know tutors are looking for. This is likely to be the case with badly constructed essays where the line-of-argument is difficult to follow. If students cannot understand the content and structure of the assignment, they may have particular difficulty understanding the tutor's formative feedback about ways to improve the work. Indeed, we may wonder whether it is feasible to separately assess some criteria (e.g. Sadler, 2008).

What are the implications for the use of exemplars? On the one hand, complete exemplars are useful because they represent 'the whole' – a completed assignment. However, students' ability to unpack the required qualities is strongly influenced by facilitative dialogue with knowledgeable others (Vygotsky 1978). On the other hand, if certain qualities are to be made visible, tutors may need to (re)construct exemplars to create clear signals to students about archetypal problems or 'good practices'. Indeed, exemplars may need to be regularly reconstructed as tutors get student feedback about their illustrative value. Another issue, particularly for written work, is length/wordcount. Are long exemplars (whether real or constructed) overwhelming to students? Would excerpts bring greater clarity even at the expense of being incomplete assignments?

As is often the case with teaching and assessment practice, there is no 'one size fits all'. Instead, all we can offer at this stage are some suggestions: that 'real' assignments may not be the best exemplars because of their inherent complexity; that *constructed* exemplars may be more effective in making assessment qualities visible; and that constructed *excerpts* may be appropriate when students are learning to 'see' criteria for the first time.

The use of exemplars in marking teams: supporting development of communities of practice

If students can benefit from dialogue around exemplars, so too can tutors in large marking teams. Whilst the *outcomes* are different – students want to create good assignments; tutors want consistency in marking – these outcomes can be facilitated through the same *process* of discussing appropriately constructed exemplars. This is the case if the underlying problem is similar: namely that assessment criteria and standards are difficult to 'explain' because an understanding of these qualities is partly tacit. For tutors marking assignments, what is important is that they interpret the criteria and standards in a consistent way, resulting in the same judgements about the quality of students' work. The fact that tutors cannot explicate these qualities does not in itself prevent them from making consistent judgements. However, without a sense of shared understanding or inter-subjectivity, consistency is unlikely.

The recognition that 'shared understanding' is critical to consistent assessment perhaps explains why the concept of 'community of practice' has gained attention in contemporary debates about HE assessment. The concept of community of practice as defined by Wenger (1998) suggests that over time, and as a result of participating in the community, members develop a shared repertoire of language and practices which facilitates shared interpretations – for example about the overall quality of student assignments. There is still scope for dissent and change, albeit within the framework of 'joint enterprise'.

Exemplars have a potentially valuable role within marking teams as mediating objects – focal points which prompt ideas, doubts, questions and suggestions to be articulated and debated. Such discussion may also be generated within marking teams

when discussing final assignments, but is that too late? Students ask for advice *before* submission of assignments, and tutors give generic guidance throughout the module. To support both tutors and (indirectly) students, exemplars may be more effective by enabling discussion before the module gets underway.

It is possible, of course, that through discussing exemplars, tutors come to realise that they hold multiple and contested interpretations of what 'quality' means. This may be the case if tutors from different disciplines come together in cross-disciplinary modules (e.g. skills-based modules in a business school). On a more positive note, at least the differences are surfaced in a way which may initiate a review of assessment policy and practice.

Conclusion

This article began by describing the problem of student (dis)engagement with formative feedback, and argued that engagement could be enhanced by time-shifting feedback so that it comes *before* final assignments are submitted. One method of achieving this is to use exemplars annotated with feedback, and this method was investigated in a large undergraduate module in a business school, using WebCT to host the exemplars and provide a discussion forum.

Student responses to the exemplar facility suggest that most found it very useful, but not necessarily for the reasons expected: their ability to see structure and layout was praised by half of the questionnaire respondents, and one-fifth added that the feedback was a useful guide to improving the quality of assignments. However, for these cohorts of students, the invitation to discuss the exemplars and feedback online was resisted: students did not want to expose their questions to public online scrutiny. The more spontaneous and informal learning space of the classroom was preferred, suggesting that the main benefits of learning through exemplars accrue through class discussion. Indeed, the lack of a quantitative effect on students' marks suggests that this instantiation of the exemplar facility was not optimal, and that more research is needed to investigate ways to develop and improve facilities such as this.

This study has shown that students are very receptive to exemplars, but that there are many questions to be considered by tutors as they design exemplar activities: for example, whether to 'construct' exemplar assignments, or use authentic student work; whether to use complete assignments or only those parts which illustrate specific criteria; and how to generate debate in order to deepen students' tacit understanding of the assessment criteria so that they develop their own skills of self-assessment.

Acknowledgements

The authors are grateful to their colleague, Berry O'Donovan, and to anonymous reviewers, who commented on an earlier draft of this paper.

Notes

1. FDTL 144/03 – HEA Project financed under the Fund for the Development of Teaching and Learning.

Notes on contributors

Karen Handley, PhD, is a principal lecturer at Oxford Brookes University in the Department of Management and Organisational Studies. Her research interests include assessment and

feedback, dialogic pedagogy, and student and staff identity. Recent projects include the FDTL5 'Engaging Students with Assessment Feedback'.

Lindsay Williams is a senior lecturer at Oxford Brookes University, in the Department of Marketing and Operations Management. Her research interests include consumer marketing and the management of plagiarism.

References

Bandura, A. 1986. *Social foundations of thought and action: A social cognitive theory.* Englewood Cliffs, NJ: Prentice-Hall.

Boud, D. 1995. *Enhancing learning through self-assessment.* London: Routledge.

Bryson, C., and L. Hand. 2007. The role of engagement in inspiring teaching and learning. *Innovations in Education and Teaching International* 44: 349–62.

Chanock, K. 2000. Comments on essays: Do students understand what tutors write? *Teaching in Higher Education* 5: 95–105.

Detterman, D., and R. Sternberg, eds. 1993. *Transfer on trial: Intelligence, cognition and instruction.* Norwood, NJ: Ablex.

Entwistle, N. 1988. Motivational factors in students' approaches to learning. In *Learning strategies and learning styles,* ed. R.R. Schmeck. New York: Plenum Press.

Gibbs, G., and C. Simpson. 2004. Conditions under which assessment supports students' learning. *Learning and Teaching in Higher Education* 1: 3–31.

Gick, M., and K. Holyoak. 1983. Schema induction and analogical transfer. *Cognitive Psychology* 15: 1–38.

Huxham, M. 2007. Fast and effective feedback: Are model answers the answer? *Assessment & Evaluation in Higher Education* 32: 601–11.

Lave, J., and E. Wenger. 1991. *Situated learning: Legitimate peripheral participation.* Cambridge: Cambridge University Press.

Millar, J., and A. Szwelnik. 2007. Warmer feedback: Engaging students with assessment feedback. Paper presented at the BMAF annual conference, May 2–4, in Oxford Brookes University, Oxford.

Moore, M. 1989. Three types of interaction. *American Journal of Distance Education* 2: 1–6.

Nicol, D. 2008. Learning is a two-way street. *Times Higher Education,* April 24.

O'Donovan, B., M. Price, and C. Rust. 2004. Know what I mean? Enhancing student understanding of assessment standards and criteria. *Teaching in Higher Education* 9: 325–35.

Orsmond, P., S. Merry, and K. Reiling. 2000. The use of exemplars and formative feedback when using student derived marking criteria in peer and self-assessment. *Assessment & Evaluation in Higher Education* 25: 23–38.

Orsmond, P., S. Merry, and K. Reiling. 2002. The use of exemplars and student derived marking criteria in peer and self-assessment. *Assessment & Evaluation in Higher Education* 27: 309–23.

Polanyi, M. 1973. *Personal knowledge.* London: Routledge/Kegan Paul.

Price, M. 2005. Assessment standards: The role of communities of practice and the scholarship of assessment. *Assessment & Evaluation in Higher Education* 30: 215–30.

Price, M., K. Handley, B. den Outer, and J. Millar. 2007. Report on case studies conducted for the FDTL5 project: Engaging students with assessment feedback. https://mw.brookes.ac.uk/display/eswaf/Case+Studies (accessed July 16, 2008).

Rust, C., M. Price, and B. O'Donovan. 2003. Improving students' learning by developing their understanding of assessment criteria and processes. *Assessment & Evaluation in Higher Education* 28: 147–65.

Sadler, D. 1989. Formative assessment and the design of instructional systems. *Instructional Science* 18: 119–44.

Sadler, R. 1987. Specifying and promulgating achievement standards. *Oxford Review of Education* 13: 191–209.

Sadler, R. 2008. Transforming holistic assessment and grading into a vehicle for complex learning, In *Assessment, learning and judgement in higher education,* ed. G. Joughin. Springer.

Salomon, G. 1992. New challenges for educational research: Studying the individual within learning environments. *Scandinavian Journal of Educational Research* 36, no. 3: 167–82.
Vygotsky, L.S. 1978. *Mind in society: The development of higher mental processes.* Cambridge: Harvard University Press.
Wenger, E. 1998. *Communities of practice: Learning, meaning and identity.* Cambridge: Cambridge University Press.