Assessment, marking and feedback: understanding the lecturers' perspective

Practitioner Research
In Higher Education

Vol 6 (2) page 3-24

Lin Norton, Bill Norton, Ian Sadler

Liverpool Hope University

nortonl@hope.ac.uk

Abstract

This study is part of a larger research project originally funded by the Write Now CETL looking at assessment, marking and feedback from the lecturers' perspective. Earlier findings have suggested that with new lecturers at least, there are some discipline differences in how able they feel they can put into practice what they have learned about assessment pedagogy on their PGCert courses. To further explore these differences with experienced lecturers, we have designed a more general tool called the Assessment, Marking and Feedback Inventory (AMFI). This was completed by 45 lecturers from a single UK university. Findings suggest that lecturers feel that one of the barriers to good assessment and feedback practice is time and workload. There were also some indications of discipline differences, with lecturers from 'hard' applied disciplines feeling this more strongly than those from the 'soft' disciplines. In a second stage of the development of the AMFI, the feedback section was refined and used in a survey of 53 academics in hospitality, leisure, sport and tourism. Findings showed broadly similar patterns with the original data, although the items related to how best to give feedback (face to face, in groups) provoked the highest rate of uncertain responses, suggesting further refinement is needed.

Keywords

Assessment methods; marking; feedback; lecturers' perceptions; discipline differences.

Introduction

In the current climate of increased tuition fees and less government funding for UK universities, students' expectations and demands are likely to increase, particularly in relation to assessment and feedback. This follows many years of a certain amount of dissatisfaction as expressed in the National Student Survey (2005–2011), although there has been a slight improvement over time. As part of the work of the Write Now CETL research team, we have been exploring assessment from the lecturers' viewpoint in a series of studies (Norton, Norton and Shannon, 2011; 2012) as a response to the need for more analysis of the lecturer perspective. In the Higher Education sector there

has been considerable pressure on lecturers to improve their assessment, marking and feedback practices. This pressure has come from governments who want more accountability, from students and parents who are often paying stakeholders, and from institutions who are anxious to preserve good ranking in league tables (derived in part from student satisfaction surveys). 'Top down' edicts, however, are unlikely to work effectively unless we have a good understanding of what lecturers themselves feel about assessment and how it fits in with the mores and epistemological foundations of their own disciplines. One size (such as assessment for learning) does not necessarily fit all. This has led to the development of the Assessment Design Inventory (ADI), intended for use as both a research tool and as a stimulus for staff development.

This earlier research is reported in full in Norton, Norton and Shannon (in revision) but one of the findings that particularly interested us was that there were differences in the subject disciplines. We used one of the best known discipline taxonomies which has been that of Biglan (1973a, b) who originally identified three dimensions to academic disciplines:

- **hard** (e.g. natural and physical sciences) versus **soft** (e.g. social sciences and humanities) disciplines;
- **pure** (e.g. mathematics, philosophy) versus **applied** (e.g. education, law) disciplines;
- **life** (e.g. agriculture) versus **nonlife** (e.g. languages) systems.

Much subsequent research appears to focus on the first two of these dimensions. An example has been the work of Smart and Elton (1982) who found that lecturers in soft disciplines (e.g. social sciences, humanities) and those whose disciplines were more concerned with practical application than theory (the applied disciplines) tended to place greater emphasis on the character development and intellectual self-actualisation of students than did their colleagues in the hard and pure disciplines.

Angelo and Cross (1993) found similar differences in that science and maths lecturers were more concerned with teaching facts and principles of their disciplines whereas lecturers in the arts were more concerned with fostering student development and personal growth. While accepting that broad frameworks such as these are themselves open to criticism (Kreber, 2009), we decided to use Biglan's typology as part of an exploratory broad brush approach to see if there were any differences between respondents on the 'hard–soft' and 'pure–applied' dimensions. Broadly speaking, our earlier results from the work on the Assessment Design Inventory showed that newly qualified lecturers from 'hard' disciplines were less likely to follow desirable assessment practices and more likely to be affected by 'constraints' than their colleagues from the 'soft' disciplines. Building on this research, we have carried out a study, using a new tool which we have been developing called the Assessment, Marking and Feedback Inventory (AMFI). The investigation, reported here, has been

carried out at a single university in the UK, described as post-1992, to see if similar differences would appear when asking more experienced lecturers about their general views on assessment, marking and feedback.

Theoretical background: the effects of the discipline

Neumann, Parry and Becher (2002) used Biglan's (1973) original classification of hard–soft and pure–applied and found differences in teaching and research. Academics from the hard subjects (both pure and applied) were strongly committed to research and saw teaching as relatively unproblematic; they also tended to favour a collaborative approach to both research and teaching. Academics in the soft (pure and applied) subjects were found to put a greater emphasis on scholarly knowledge that translated into teaching; they also put more emphasis on individualistic enquiry rather than on joint teaching.

Since assessment is so closely aligned to learning and teaching, it seems reasonable to assume that there would be similar differences in lecturers' assessment, marking and feedback practices and attitudes, related to their subject disciplines**.** However, we know of relatively little research in this area**.** An exception has been the work of White and Liccardi (2006), who drew some distinctions between the disciplines and assessment methods favoured. They described lecturers in hard pure subjects (e.g. natural sciences) favouring assessment methods that reflected their view of the quantitative nature of knowledge, such as exam-based assessment with specific and focused exam questions, and objective tests. Lecturers in hard applied subjects (e.g. engineering) preferred exam questions about problem solving. Lecturers in soft pure subjects (e.g. social sciences and humanities) favoured essay questions, short answer questions, oral presentations and continuous assessment. Finally, lecturers from soft applied subjects (e.g. nursing and education) also favoured essays, but in addition preferred project-based assignments and tended to encourage peer- and self-assessments.

It was to explore these differences further that the present study was designed. We were keen to use Shulman's (2005) concept of signature pedagogies which he defined as:

> *'…a mode of teaching that has become inextricably identified with preparing people for a particular profession.'*

Shulman carried out a 10-year study to understand how people are prepared for practice in law, engineering, the clergy, teaching, nursing and medicine. He argued that education for professionals has to include more than knowing, and that understanding the discipline is only part of what we should be teaching. Shulman described a mode of teaching that has become inextricably identified with preparing people for a particular profession; he said there were three defining characteristics:

> *'…it's distinctive in that profession…*

> *'…it is pervasive within the curriculum. So that students learn that as they go from course to course, there are certain continuities that thread through the program that are part of what it means to learn to "think like a lawyer", or "think like a physician", or "think like a priest."*
>
> *'(it)…cuts across institutions and not only courses. Signature pedagogies have become essential to general pedagogy of an entire profession, as elements of instruction and of socialization.'*

Interestingly, he also clearly stated that signature pedagogies are not stable and unchanging and are as likely to change as professional practices themselves. To this we would add academic disciplines which also are subject to fluctuation. In so doing, we acknowledge Poole's (2009) point that professions are not the same as disciplines, they contain disciplines and so professional pedagogy must of necessity be created across disciplines. Since Shulmans' original work, Gurung et al. (2009) discuss how different disciplines have approached this concept by describing it as teaching disciplinary habits of mind. They examine how such 'habits of mind' work in the fields of humanities, fine arts, social sciences, natural sciences and maths. Their work has been further updated and extended (Chick et al., 2012) by focusing more on interdisciplinary fields and programmes.

Although Shulman wrote about teaching processes, he appears to be silent on the subject of assessment processes. Taking the very broad concept of signature pedagogies as a proxy for teaching process differences, then there presumably must be assessment process differences. It was this issue that we wanted to pursue in relation to lecturers' perceptions of using different assessment methods, and their practices and attitudes to marking and feedback. This assumption is given some support by the work of Reimann (2009) who used the term 'ways of thinking and practising' (WTP) to point out the challenges for assessment by citing an academic developer who said:

> *'…there is in many of these accounts a realization that people who would be otherwise regarded as equal on conventional methods of assessment in fact differ quite markedly…to think like an economist or whatever'* (Reimann, 2009:93).

Research aims

We were specifically interested in two main questions:

1. What do lecturers think about assessment methods, marking and feedback practices?
2. Are there any differences between the subject disciplines?

Questionnaire development

Stage 1

A pilot version of the AMFI was developed which consisted of four main sections. The items were derived from over 80 interviews with lecturers from five UK institutions. This initially produced over 600 items using N-Vivo. These were progressively refined by iterative readings from a team of four researchers who met regularly to agree a more manageable number. One of our guiding principles was that there should be sufficient items to capture what we wanted to measure but not so many that participants would be unwilling to complete such a questionnaire. At this stage it was decided to separate items relating to assessment design into one inventory (ADI) leaving the other items relating to types of assessment, marking and feedback to be developed into a second inventory (AMFI). Further work was then carried out by a post-doctoral member of the research team who consulted the assessment literature and produced the version described here. This resulted in a fairly lengthy inventory which consisted of four main sections:

1. methods of assessment (52-item checklist from Bloxham and Boyd (2007)
2. rationale for choice of assessment method (13 statements)
3. marking attitudes and practices (22 statements)
4. feedback attitudes and practices (31 statements).

Respondents were also asked in a free text box to add any other comments they might have related to assessment marking and feedback, as well as any comments about the AMFI questionnaire design.

Stage 2

In a second stage of development, the third author used a refined version of the section on feedback (now reduced to 17 items) as part of a larger survey for the HEA on academics' attitudes towards feedback and case examples of feedback practice in the subject areas of hospitality, leisure, sport and tourism (HLST) (Sadler 2011).

Procedure

Ethical clearance for the research was obtained from the university, following the institutional ethical requirements. In stage 1, an email was sent round to all the lecturers in the university inviting them to take part in the study by completing an online version of the AMFI. A modest prize draw of three £50 Amazon vouchers was offered to encourage participation. Out of the 162 lecturers, 45 from 17 disciplines completed the AMFI. This represented a 28% response rate, which is just below the average response rate for online surveys as cited by the Instructional Assessment

Resources (IAR) website. This was judged acceptable for an exploratory study but to be cautious when using inferential statistics, only non-parametric tests were used (Siegel, 1956). In stage 2, which related only to the feedback section of the AMFI, the survey was circulated electronically to academics using a 'HLST key contacts' mailing list. Participants were informed that the data might be used for validation purposes. A total of 67 participants from 25 teaching establishments took part but not all completed the feedback section, which left us with a final total of 53 responses to analyse.

Representation of the subject disciplines

One of the main research aims was to analyse the AMFI results to see if there were any subject differences. Using Biglan's (1973) original classification, later refined by Stoecker (1993) and Nelson Laird et al. (2008), we divided our participants into their representative subjects as hard (pure and applied) and soft (pure and applied). Since not all respondents stated their subject, this left us with 41 as shown in Table 1. In stage 2, although all four subject areas (hospitality, leisure, sport and tourism) were represented, we did not attempt to classify participants using Biglan's categories as the majority would be described as 'soft applied'.

Table 1. Classification of AMFI respondents according to their subject disciplines in stage 1.

Categories (Biglan, 1973, Stoecker, 1993, Nelson Laird, 2008).	**Subjects**
Soft applied (N=24)	Arts Business studies Childhood and youth studies Disability studies Drama and theatre studies Education Marketing Media Sport
Soft pure (N=12)	English Geography Music Psychology Sociology Theology and religious studies
Hard applied (N=5)	Computing Health
Hard pure (N=0)	-

Results

What methods of assessment are currently being used?

The first element we looked at was the responses to the different types of assessment method. Fifty-two methods as delineated by Bloxham and Boyd (2007) were presented in the first section of the AMFI and participants were asked to look at each one and indicate those methods and respond:

- 'Have used'
- 'Would like to use'
- 'Not familiar'

The results of this analysis are presented in the form of percentage responses for each method in Table 2. Since this initial analysis did not look for discipline differences, the full number of 45 participants was used.

Table 2. Frequency of responses to methods of assessment (Bloxham and Boyd, 2007) by percentage (N=45).

Method	**Have used %**	**Would like to use %**	**Not familiar %**
Essay	98	0	0
Dissertation	93	2	0
Research project	93	0	0
Presentation	91	0	0
Reflective learning assignments	80	2	4
In-class tests	76	9	0
Examinations (unseen, seen, open-book, case study, take-away)	75	7	4
Portfolio (written)	71	9	2
Evaluation of journal article or other paper	69	18	9
Problems and case study analysis	66	11	7
Review of book, article, website, etc.	66	20	2
Project	62	7	7
Students-led seminar or discussion	53	16	7
Poster	51	13	11
Tests	51	7	11
Direct observation	49	16	11
Multiple choice tests	49	16	11

Method	**Have used %**	**Would like to use %**	**Not familiar %**
Exhibition and displays	47	18	13
Data Interpretation Exercise	40	16	29
Short-answer questions	40	16	16
Debate speech	37	29	16
Viva voce examination	37	18	13
Computer based assignment	36	22	24
Fieldwork reports	35	20	24
Writing abstracts	35	31	11
Laboratory examination and practical tests	33	4	22
Professional tasks	33	11	29
Placement reports	29	20	24
Annotated bibliography	27	22	33
Modified essay questions	27	9	33
Work tasks: newspaper articles, press release, executive summaries	26	9	29
Completion exercise	24	4	60
Simulation exercises	24	24	22
Design task (including manufacture)	22	7	49
Electronic presentation (web pages)	22	38	13
Performance and production	22	11	36
Portfolios and sketch books	20	7	31
'Crits'	18	0	67
Work books	18	7	31
Laboratory reports	16	7	36
Lay commentary on specialist material(e.g. journal article)	16	18	38
Concept map	11	20	49
Exchange or sandwich year reports	9	16	53
Film or radio programmes	9	29	36
Grant applications	9	13	40
Geological mapping	4	7	56
Marking glossaries	4	16	51
Capstone assignment	2	7	76
Objective structured clinical examination	2	9	53

Method	**Have used %**	**Would like to use %**	**Not familiar %**
Synoptic examination	2	11	49
Treatment reports	2	7	51
Patchwork texts	0	7	62

Of the 52 methods of assessment presented, it was not surprising to see that the most popular methods (70% and over) were essays, dissertations, research projects, presentations, reflective learning assignments, in-class tests, examinations (of all kinds) and portfolios. It was interesting to see that only one assessment method (patchwork texts) had not been used by any of the respondents, although three lecturers said they would like to use this method. Nineteen of the 52 methods were unfamiliar to more than a third of our lecturers. Finally, many respondents expressed interest in using methods of assessment they did not currently use. Two of the respondents suggested in the free text box that an additional response category of *'not appropriate to my subject'* would have been helpful, as they were obliged to not respond to some methods as they did not have this option.

Looking next at the discipline differences with the 41 participants who had identified their subjects showed some small, but not statistically significant, differences in the most commonly used assessment methods. All three disciplines had 100% response rates for essays, research projects, and presentations, but lecturers from the soft subjects (both pure and applied) did not all use reflective learning assignments or written portfolios, which was a fairly surprising finding. In addition, not all soft applied lecturers used dissertations, in-class tests or exams. Since the actual numbers are so small, it would be unwise to draw any conclusions here, but there is enough indication to suggest further research would be useful to support White and Liccardi's (2006) broad findings.

Overall, the findings suggest a wide array of methods were being used by our participants, although traditional methods were still the most favoured. As Boud (1990) says, however, it is not the method that is so important but the purpose of it and how that articulates with the overall purpose of the course or module. This led to the design of the second section of the AMFI.

What is the rationale for the choice of assessment methods?

This section consisted of 13 statements but for this analysis four items were excluded due to floor/ceiling effects (A floor effect means that all the participants responded that the item was completely unimportant; a ceiling effect means that all the participants responded that the item was very important). A further item was excluded

as the wording was almost identical to an existing item. The descriptive means for the remaining eight items are presented in Table 3, where the following scoring system was used:

'Completely important' = 4

'Important' = 3

'Not important' = 2

'Completely unimportant' = 1

Table 3. Assessment methods rationale: discipline differences in means

The assessment method chosen should take into account whether or not it…	**Hard applied (N=5)**	**Soft pure (N=12)**	**Soft applied (N=24)**	**Overall (N=41)**
lessens the likelihood of students cheating	2.80	3.17	3.00	3.02
incorporates both essays and exams	2.60	3.08	2.42	2.63
creates memory stress	2.00	2.08	2.04	2.05
allows students to develop oral as well as written skills	3.00	3.33	3.42	3.34
allows for a variety of exam formats	2.40	3.00	2.92	2.88
allows for a variety of word lengths in written assignments	2.80	2.83	2.96	2.90
allows for group work	2.80	3.08	3.00	3.00
takes into account students' preferences	2.80	2.83	2.75	2.78

In this analysis, there were no statistically different discipline differences, so the findings will be discussed for the whole sample. Looking at the overall means, we can see that the two most important rationales are to allow students the opportunity to develop oral as well as written skills and to take part in group work. This would seem to indicate a rationale for developing skills perhaps linked to what Knight and Yorke (2004) have called the USEM (understanding, skillful practices, efficacy beliefs and metacognition) but as Yorke argues, it applies more to student success in general. The third most important rationale for choosing an assessment method is to lessen the likelihood of students cheating. Preventing students from plagiarising is not a pedagogical rationale, but is commonly cited for using specific methods of assessment such as portfolios (Irons, 2004) and exams (Norton et al., 2006).

Marking: practices and attitudes

This section consisted of 22 statements in which practices and attitudes were mixed together to minimise response set (such as always ticking the same response without considering the meaning of the item, or ticking an item to give a 'socially desirable response). Participants were asked to rate how much they agreed or disagreed with each statement using the scale of:

'Strongly agree' = 4

'Agree' = 3

'Disagree' = 2

'Strongly disagree' = 1

As before, items producing floor or ceiling effects were discarded from the analysis. This left a total of 4 items reflecting practices and 14 reflecting attitudes. These have been separated out for ease of reading in Table 4.

Table 4. Marking practices and attitudes: discipline differences in means.

Marking practices	**Hard applied (N=5)**	**Soft pure (N=12)**	**Soft applied (N=24)**	**Overall (N=41)**
I have learned to mark myself without any professional training	3.25	2.83	2.68	2.79
I divide the marks according to the weight of each assessment criterion	2.80	2.22	2.62	2.54
I give more marks to the theoretical background of an assignment	2.80	2.40	2.68	2.62
I advise students not to rely too heavily on the mark scheme when writing their assignments	2.60	2.00	2.04	2.10
Marking attitudes	**Hard applied (N=5)**	**Soft pure (N=12)**	**Soft applied (N=24)**	**Overall (N=41)**
The external examiner system does not ensure grading fairness	2.00	2.42	2.53	2.43
External examiners should not be given the authority to determine grades awarded to students	3.00	2.50	2.86	2.76
Grading student assignments is open to subjective interpretation	3.20	2.92	3.00	3.00
Professional training in marking will improve lecturers' marking skills	2.40	3.08	2.87	2.88
Professional training in marking makes no difference to lecturers' marking skills	2.60**	1.83	2.14	2.10
Certification/grading shows that students have achieved a certain level	3.00	3.08	2.78	2.90
Highly detailed guidelines increase the likelihood of students challenging their grades	3.40*	2.27	2.09	2.31
Accuracy in marking is part of the external examiners' responsibility	2.20	2.36	2.45	2.39
Marking is not an objective empirical process	2.40	2.75	2.79	2.72
Some students achieve high marks by luck	2.75	2.33	2.26	2.33
Marking is not an exact science	3.00	3.00	3.05	3.03
Marking is a lecturer's own professional judgement	3.40	3.08	2.91	3.03
Marking procedures are too bureaucratised	3.20	2.33	2.33	2.45

* $p<0.05$ **$p<0.01$

Looking at the overall mean scores for the four practices first of all, it can be seen that the strongest agreement was with the item that said they had learned to mark themselves without any professional training. There was a fairly strong agreement about marking for theoretical background and slightly less agreement for dividing marks according to assessment criteria. The practice that attracted least agreement, but still over a mean of 2.0, was about advising students not to rely too heavily on marks schemes when working on their assignments. This suggests perhaps some conflict since their previous response indicates they do themselves use such marking schemes. A Kruskal–Wallis one-way analysis of variance was then used to see if there were any disciplinary differences on any of the marking practices, but none was found.

When it came to attitudes to marking, we again looked first at the overall mean scores and found that lecturers tended to agree that marking was a subjective process which relies on their own professional judgment rather than on being an exact science. Two significant disciplinary differences were found. Firstly, lecturers from hard applied disciplines scored higher than those from soft pure (but not soft applied) disciplines in their degree of agreement that '*Professional training in marking makes no difference to lecturers' marking skills*'. This was statistically significant beyond the 0.01 level. Secondly, lecturers from hard applied disciplines scored higher than those from both soft pure and soft applied disciplines in their degree of agreement that '*Highly detailed guidelines increase the likelihood of students challenging their grades.*' This was statistically significant beyond the 0.05 level.

The open-ended comments about marking were around the subjectivity/objectivity debate:

> *The questions are restrictive, for example, on whether marking is subjective or objective. In my opinion marking does entail an element of subjectivity but lecturers should strive for objectivity.*

This is an interesting statement as it shows the tension that lecturers feel when exhorted to be fair and transparent, which implies an objectivity that is often difficult to achieve (Shay, 2005), despite the Quality Assurance Agency (2006) code of practice on assessment of students in which it is stated that:

> *'Institutions publicise and implement principles and procedures for, and processes of, assessment that are explicit, valid and reliable'* (QAA, 2006:8).

Interestingly, the other two comments related to marking show the effect of being overly prescriptive but this is seen as a local (i.e. the department and the university) requirement rather than one derived from the QAA.

> *My overall feeling is that we 'over mark' as opposed to over assess, which I think we also do. I think that experienced lecturers should be trusted to get marking right through, as Sue Bloxham puts it, 'connoisseurship'. In my view some of the marking schemes we use [in the xxx department] are far too*

> *detailed/trivial and discourage students from being creative or put them off exploring – they become focused on 'getting it right'. Some of the marking schemes I have had to use are absurd in their detail/prescription.*
>
> *…it is very noticeable that since the university has requested that assessment/marking criteria coupled with displaying every assessment component out of 100, students have become more fixated with chasing markings than actual learning in some cases and question like 'why did I only get 6/10 for criteria xxx when I followed the criteria...' or 'you gave me positive feedback on xxx but only gave me 7/10 surely it was worth more because I got positive feedback...' or 'where is the marking scheme?' is very worrying. The university assessment processes do not seem to be encouraging independent thinkers, creators or learners.*

It may be that lecturers 'on the ground' are not always aware of the need for institutional compliance. In the QAA code of practice, for example, assessment (or marking, or grading) criteria are mentioned 19 times. It would be difficult for universities to ignore such a strong message.

Overall, our findings suggest that there was a general tendency to believe that marking was more objective than subjective. While there was no discipline difference in marking practices, there may be some slight difference in attitudes between lecturers from hard applied disciplines and their colleagues from the soft disciplines, but the fact that this only applied to two items together with the small sample size, means we cannot draw any firm conclusions here. What was interesting in this section, though, was the open-ended comments, which although not showing disciplinary differences did show the pressures that some of our sample were feeling, and they felt these were due more to local managerial requirement than a consequence of the UK quality assurance systems.

Feedback: practices and attitudes (stage 1)

This section of the AMFI consisted of 31 items. Again, practice and attitude items were mixed together in the questionnaire itself but have been presented separately in the results as shown in Table 5. The same scale of response was used (i.e. strongly agree to strongly disagree; 4–1). For the analysis, eight items were excluded which left six practice items and 17 attitude items. These were analysed for discipline differences using Kruskal–Wallis analysis of variance.

Table 5. Feedback practices and attitudes: discipline differences in means.

Feedback practices	**Hard applied (N=5)**	**Soft pure (N=12)**	**Soft applied (N=24)**	**Overall (N=41)**
I have learned about giving feedback through professional training	2.20	2.67	2.27	2.38
I have learned about giving feedback from my colleagues	3.20	3.17	3.00	3.08
I seek advice from colleagues for providing clear information to students	2.80	2.75	2.71	2.74
I prefer to informally discuss the feedback with students	3.00	2.75	2.86	2.84
It is constructive for me and students to meet them individually at the beginning of the term	2.60	2.75	3.00	2.87
Lack of time due to workload prevents me giving good quality feedback	3.40*	2.08	2.64*	2.56
Face to face communication with students is the best way to give feedback	3.00	2.91	3.18	3.08
Feedback attitudes	**Hard applied (N=5)**	**Soft pure (N=12)**	**Soft applied (N=24)**	**Overall (N=41)**
Verbal feedback is more useful than written feedback	2.60	2.17	2.40	2.35
Students' learning depends upon the lecturer's feedback	2.60	2.58	2.56	2.57
Feedback should be about helping students understand assessment criteria	2.20	2.75	2.81	2.71
When giving feedback, it is more important to make negative rather than positive comments	1.80	1.75	1.55	1.64
Knowing students personally helps lecturers give constructive feedback	2.60	3.17	3.05	3.03
It is good to give feedback in the middle of a class in order for other students to hear	2.60	1.64	2.29	2.14
Giving feedback puts an unfair workload on lecturers	2.40*	1.75	2.14*	2.05
The requirement to give good quality feedback to a large number of students affects the quality of teaching and students' learning	2.40	2.33	2.64	2.51
Regular negative feedback can indicate that students are not responding to the teaching strategy and that an alternative strategy should be sought	2.60	2.58	2.65	2.62
Students do not care about lecturers' feedback	2.20	2.08	2.18	2.15

Looking first at the overall means for the six practices we see that the only item with an overall mean over 3.0 was the one to do with learning about giving feedback from colleagues. This relates to the earlier finding that lecturers tend to mark themselves without any professional training. There was only one feedback practice item which showed a significant disciplinary difference. Lecturers from hard applied disciplines scored higher than those from both soft pure and soft applied disciplines in their degree of agreement that '*Lack of time due to workload prevents me giving good quality feedback*'. This was significant beyond the 0.05 level.

There was a further difference with lecturers from soft applied disciplines scoring higher than those from soft pure disciplines, which was also beyond the 0.05 level of significance. Put simply, lecturers from hard applied disciplines were more likely than those from both soft pure and soft applied disciplines to feel that lack of time prevented them giving good quality feedback and lecturers from soft applied disciplines felt this more strongly than those from soft pure disciplines.

When looking at the attitudes to feedback, overall means suggested that lecturers preferred face to face feedback and knowing students personally. They also preferred to make feedback follow the positive, negative positive formula, which was echoed in the low mean agreement for making negative rather than positive comments. When analysing the attitudes for discipline differences, there was further evidence that time and workload were seen as hindering good feedback, with lecturers from hard applied and soft applied disciplines scoring higher than those from soft pure disciplines in their degree of agreement that '*Giving feedback puts an unfair workload on lecturers'* which was significant beyond the 0.05 level.

The comments from the free text box add some further illumination to this issue of time and workload, with four lecturers specifically commenting on it, with varying degrees of frustration:

> *Time is an issue and should be in the workload model*
>
> *I had 56 students in one of my groups and was the sole lecturer. This made individual personal feedback virtually impossible apart from those who e-mailed me separately (which they were all invited to do) and/or made appointments to come and see me...Staff who have high teaching loads, including high numbers in each group they teach, have major difficulty feeding back effectively to students and monitoring the effectiveness of the feedback*
>
> *I think that the measured kind of response I wanted to convey was offset by the direct nature of these statements – a product of most surveys. In general my thoughts are that it is best to give feedback within approximately three weeks, that it should be typed, constructive – in so far as this is possible – and provide pointers for future work. Moreover, best practice is for the tutor to sit down with the individual student and go through the comments. Having said all this, there*

are times when 1) a student needs telling straight if a piece of work is garbage, unacceptable; 2) it may not be possible with large groups given the constraints of time and workload – likely to get worse not better in coming years

Time can be an issue in preparing and giving feedback, especially if students are out on a placement and do not come into contact with their tutor. Finding a way to provide helpful feedback at a distance is a major issue for me, especially as written work cannot always be returned if students are on placement and also being aware that students tend to look at the mark rather than the comments on how to improve. There needs to be an awareness that not all assessment is formal. Whilst the main form of academic assessment is through assignment students on my course are also informally assessed through their contributions to group presentations, seminar participation – some form of feedback needs to be given on these to highlight other qualities not recognised through assignments. Lesson observation is also used to assess student progress against national teaching standards.

This last comment is particularly interesting as it highlights one reason why discipline differences are important. In this case (teacher education), feedback takes on many aspects including that of professional practice.

The only other statistically significant difference in attitudes was in the item about group feedback where lecturers from hard applied and soft applied disciplines scored higher than those from soft pure disciplines in their degree of agreement that '*Feedback sessions in groups are more effective than individual feedback'*. This was significant beyond the 0.05 level.

Overall, our analyses on feedback from stage 1 show two main findings. Firstly, there seemed to be little evidence of professional training in giving feedback and that there was a general consensus about making feedback as personal as possible. Secondly, there was an issue with time impacting negatively on feedback shown in the questionnaire analysis and in the free text comments. In our results, the lecturers from the hard applied subjects seemed to feel this more strongly, although the comments showed that professionally orientated disciplines such as teacher education contextualise the whole notion of feedback differently.

Stage 2

In developing the section on feedback for the HSLT survey, cluster analyses and further refinement was carried out which left us with a total of 17 items, some of which were considerably revised (see Table 6), and the response scale was changed with the additional option of *'uncertain'* (5) to further judge the usefulness of the items. Table 6. shows a comparison of the stage 1 data overall means with that of the stage 2 HSLT data overall means, having eliminated any uncertain responses in order to compare like with like.

Table 6. Feedback attitudes: HLST (stage 2) study frequencies/means and comparison with university (stage 1) study means.

Feedback attitudes *Bracketed italics denote wording used in stage 1 study*	**HLST Agree %**	**HLST Disagree %**	**HLST study Mean (*N*)***	**University study Mean (N=41)**
Lecturers learn best *(I have learned)* about giving feedback through professional training	42	45	2.52 (*46*)	2.38
Lecturers should *(I)* seek advice from colleagues for providing clear information to Students	77	9	3.04 (*46*)	2.74
Lecturers *(I)* prefer to informally discuss the feedback with students	28	34	2.42 (*33*)	2.84
Lack of time due to workload prevents lecturers *(me)* giving good quality feedback	60	38	2.73 (*52*)	2.56
Face to face communication with students is the best way to give feedback	81	8	3.27 (*47*)	3.08
Verbal feedback is more useful than written feedback	42	40	2.60 (*43*)	2.35
Students' learning depends upon the lecturer's feedback	79	15	3.06 (*50*)	2.57
Feedback should be about helping students understand assessment criteria	51	42	2.62 (*49*)	2.71
Knowing students personally helps lecturers give effective *(constructive)* feedback	64	26	2.96 (*48*)	3.03
It is good to give feedback in the middle of a class in order for other students to hear	34	47	2.30 (*43*)	2.14

Looking at this table we can see that there is a consistency in responses between the two samples with only minor variations in means. What is also important to note is that there were two items which generated a high response rate of '*uncertain'*. These were: '*Lecturers prefer to informally discuss the feedback with students'* and '*Feedback sessions in groups are more effective than individual feedback*'. The latter item had thrown up some discipline differences in stage 1, so it is important that in further refinements of the AMFI we consider whether some rewording is needed.

Conclusions

In the first phase of this study we were looking to explore experienced lecturers' views of assessment methods, marking and feedback. Given that the main analysis using the whole AMFI was from a small sample from one UK institution, it would be unwise to generalise from these results, but we would want to suggest that the findings indicate that this is an area worth exploring more widely. In particular, the lack of professional training in both marking and giving feedback would seem to raise an issue that academic developers might wish to explore. This is not to necessarily claim that the only way of helping academics learn how to mark is through explicit training but that as Jawitz, (2008; 2009) suggests there should be '*opportunities for conversations about assessment practice*', as tacit knowledge about standards of marking can be difficult to articulate. Our findings also suggest that practical barriers such as lack of time, high student numbers and heavy workloads militate against good marking and feedback practices.

The institutional context of this research is an important one and may in some part account for the findings in stage 1, which was carried out in a single university. In some earlier research we found differences between institutions in terms of assessment design (Norton et al., 2011, in submission). More generally, Elen et al. (2007:123) found that in traditional universities teaching is more research-centred than student-centred and that:

> *'…the link between research and teaching is fundamentally based on and directed towards a mature epistemological disposition.'*

They go on to argue that these two features should be explicitly considered in staff development processes. Holligan et al. (2011) also argue that how the same subject is taught and researched is often determined by the type of university in which it is offered. Thus we are cautious in making too much of our findings in terms of implications for staff development, but the findings from stage 2 also suggested there was a low level of agreement (42%) that professional training about giving feedback was desirable, and here the data was collected from a wide range of teaching institutions.

Finally, it was regrettable that our sample in stage 1 did not include lecturers from the hard pure discipline category and so few from the hard applied category. However,

there were enough indications to suggest that there may be some disciplinary differences with 'hard' lecturers agreeing more strongly than the 'soft' lecturers that professional training was not important and that students are more likely to challenge grades if they are given too much detail about how they are marked. They also agreed more strongly that they were adversely affected by lack of time and heavy workloads which might account for their preference for giving feedback in a group than individually. While we do not want to make too much of it at this point, these preliminary results would appear to support Neumann, Parry and Becher's (2002) observation that lecturers from hard disciplines tended to see teaching (and thereby assessment) as relatively unproblematic. They also appear to support our earlier findings still to be published related to assessment design (Norton, Norton and Shannon, in revision) which indicate that lecturers from hard disciplines were more likely to indicate more constraints to 'desirable' assessment design practice, than their colleagues from the soft disciplines. This is perhaps some way from Shulman's concept of signature pedagogies, but the discipline perspective is one that we feel is worth pursuing.

Development of the AMFI

The AMFI has been further refined and we plan to use it in a number of universities both within and outside the UK. Anyone wishing to become involved in this research is welcome to contact the first author for more details.

Acknowledgments

The initial version of the AMFI was developed by Kamel Mansi, a former post-doctoral researcher at Liverpool Hope University for the Write Now CETL, which funded this research.

References

All websites accessed 2 November 2012

Angelo, T. , Cross, P. (1993)*Classroom assessment techniques*. 2nd ed. San Francisco: Jossey- Bass

Biglan, A. (1973a). Relationships between Subject Matter Characteristics and the Structure and Output of University Departments. *Journal of Applied Psychology,* 57 (3),204 - 213.

Biglan, A. (1973b) The characteristics of subject matter in academic areas. *Journal of Applied Psychology* 57, 195–203.

Bloxham, S., Boyd, P. (2007) *Developing Effective Assessment in Higher Education.* Open University Press.

Boud, D. J. (1990) Assessment and the promotion of academic values. *Studies in Higher Education* 15(1): 101–111.

Chick, N.L., Haynie, A., Gurung, R.A.R. (2012) *Exploring More Signature Pedagogies: approaches to teaching disciplinary habits of mind.* Virginia: Stylus.

Elen, J., Lindblom-Ylanne, S., Clement, M. (2007) Faculty development in research-intensive universities: The role of academics' conceptions on the relationship between research and teaching. *International Journal for Academic Development* 12(2): 123–139.

Gurung, R.A.R., Chick, N.L., Haynie, A. (2009) *Exploring Signature Pedagogies: approaches to teaching disciplinary habits of mind.* Virginia: Stylus.

HEFCE (2011) 2011 National Student Survey Summary Data. http://webarchive.nationalarchives.gov.uk/20100202100434/http://www.hefce.ac.uk/learning/nss/data/2011/

Holligan, C., Wilson, M., Humes, W. (2011) Research cultures in English and Scottish university education departments: an exploratory study of academic staff perceptions. *British Education Research Journal* 37(4): 713–734.

Irons, A. D. (2004*) Using portfolios in assessment to reduce plagiarism*. http://plagiarismadvice.com/documents/papers/2004Papers12.pdf

Jawitz, J. (2008) Learning to assess in the academic workplace: case study in the natural sciences. *South African Journal of Higher Education* 22(5): 1006–1018.

Jawitz, J. (2009) Learning in the academic workplace: the harmonization of the collective and individual habitus. *Studies in Higher Education* 34 (6): 601–614.

Kreber, C. (2009) (ed) *The University and its Disciplines. Teaching and learning within and beyond disciplinary boundaries.* London: Routledge.

Knight P.T., Yorke, M. (2004) *Learning, Curriculum and Employability.* London: RoutledgeFalmer

Nelson Laird, T.F., Shoup, R., Kuh, G.D., Schwarz, M.J. (2008) The effects of discipline on deep approaches to student learning and college outcomes. *Research in Higher Education* 49(6): 469–494.

Neumann, R., Parry, S., Becher, T. (2002) Teaching and learning in their disciplinary contexts: A conceptual analysis. *Studies in Higher Education* 27(4): 405–417.

Norton, L., Harrington, K., Norton, B., Shannon, L. (2006) Challenging traditional forms of assessment: university teachers' views on examinations. *Making a Greater Difference: Connecting to Transformational Agendas.* Conference of the International Society for the Scholarship of Teaching and Learning (ISSOTL), Washington DC, USA, 9–12 November. University of Wisconsin Colleges.

Norton, L., Norton, B., Shannon, L. (in revision) Revitalising assessment design: what is holding new lecturers back?

Norton, L.S, Norton, B., Shannon, L. (2012) The impact of UK university teaching programs on lecturers' assessment practice: a case for pedagogical action research. In E. Simon and G.Pleschová (Eds), *Teacher development in higher education. Existing programs, program impact, and future trends,* Abingdon: Routledge. Chapter 11: 191-210.

Norton, L., Norton, B., Shannon, L. (2011) The Assessment Design Inventory: A tool for research and staff development. In Rust (Ed.) *Improving Student Learning: Improving Student Learning Global Theories and Local Practices: Institutional, Disciplinary and Cultural Variations.* ISL 18. Proceedings of the 2010 ISSOTL/ISL conference. Oxford: Oxford Centre for Staff and Learning Development. Chapter 4 section 20: 199–214.

Poole, G. (2009) Academic Disciplines: Homes or Barricades? in Kreber, C. (ed)*The University and its Disciplines. Teaching and learning within and beyond disciplinary boundaries*, London: Routledge, 50–57.

Quality Assurance Agency (QAA) (2006) *Code of practice for the assurance of academic quality and standards in higher education.* Section 6. Assessment of students. http://www.qaa.ac.uk/Publications/InformationAndGuidance/Documents/COP_AOS.pdf

Reimann, N. (2009) Exploring disciplinarily in academic development: Do 'ways of thinking and practising' help faculty to think about learning and teaching? In Kreber, C. (ed) *The University and its Disciplines. Teaching and learning within and beyond disciplinary boundaries*, London: Routledge, 84–95.

Sadler, I. (2011) Hospitality, Leisure, Sport and Tourism (HLST). A review of contemporary practice: Feedback. *Higher Education Academy* http://www.heacademy.ac.uk/assets/hlst/documents/critical-reviews/FeedbackHLST.pdf

Shay, S. (2005) The assessment of complex tasks; a double reading. *Studies in Higher Education* 30(6): 663–679.

Shulman, L. (2005) Signature pedagogies in the professions. *Daedalus* 134(3): 52–59. *MIT Press Journals* http://www.mitpressjournals.org/doi/abs/10.1162/0011526054622015

Siegel, S. (1956). *Nonparametric Statistics for the Behavioural Sciences*. New York: McGraw-Hill.

Smart,J.C., Elton,C.F. (1982) Validation of the Biglan Model. Research inHigher Education 17(3):213-229

Stoecker, J. L. (1993) The Biglan classification revisited. *Research in Higher Education* 34(4): 451–464.

The University of Texas at Austin (2011) *Instructional Assessment Resources (IAR).* http://www.utexas.edu/academic/ctl/assessment/iar/teaching/gather/method/survey-Response.php

White, S., Liccardi, I. (2006) Harnessing insight into disciplinary differences to refine e-learning design. Frontiers In Education (FIE) conference, San Diego, California, 28-31 October. IEEE Xplore Digital Library. http://ieeexplore.ieee.org/xpl/login.jsp?tp=&arnumber=4117208&url=http%3A%2F%2Fieeexplore.ieee.org%2Fstamp%2Fstamp.jsp%3Ftp%3D%26arnumber%3D4117208