

Authenticity in assessment tasks: a heuristic exploration of academics' perceptions

Effie Maclellan*

University of Strathclyde, Glasgow, UK

This study was an exploration of the extent to which higher education tutors' perceptions of assessment were consistent with the construct of authenticity. Depth interviews with 12 academics sought views on what might constitute desirable assessment tasks and scoring methods to use with students. Summaries of transcribed interviews suggested that assessment should focus on real world problems and have some meaning to a real world audience. However, analysis of the interviews using Newmann's (1997) criteria for authentic assessment showed that, collectively, interviewees' conceptualisation of authenticity was incomplete. The results are discussed in terms of constructivist perspectives on learning.

The issue

Three different strands of evidence coalesce to give rise to the issue of concern, which drives this paper. Firstly, while professionally oriented/vocational courses in higher education are rightly concerned with the constituents of student competence on field experience, the in-faculty components of such courses perhaps do not always promote the critical thinking and reflection which enables autonomous professional development to continue (Weil & McGill, 1989; Eraut, 1994). This imbalance is reflected in the emphasis placed on 'professional performance' by specialised accrediting bodies and without detracting from the significance of placement experience in students' learning, there is nevertheless a persuasive argument that it is the in-faculty components of professional preparation which can provide a supportive, developmental and coherent curriculum based on a clear conceptualisation of the professional role (Bowden & Marton, 1998; Roth, 1999). The teaching/assessment of in-faculty components is therefore worthy of closer examination.

Secondly, assessment has a significant impact on students' approaches to learning (Ramsden, 1997; Gibbs, 1999). If students perceive a need to understand the material in order to successfully negotiate the assessment task, they will engage in deep learning but if they perceive the assessment instrument to require rote learning

of information, they will be unlikely to engage with the higher level objectives which may well have been intended by the programme of study (Crooks, 1988; Gibbs, 1999). Given this, it seems that assessment tasks themselves are critically important.

Thirdly, it appears that one's epistemological beliefs are related to one's behaviour. Thus, how academics teach is fundamentally influenced by how they conceptualise the task of teaching (Kember, 1997, 1998; Pajares, 1992; Prosser & Trigwell, 1999; Martin *et al.* 2000; Hativa, 2000). From this it follows that academics' conceptions of assessment may well affect their assessment practices.

In summary, given the pedagogical importance of in-faculty professional preparation, the influence of assessment tasks on the quality of student learning, and the influence of academics' conceptualisation of assessment on their practice, it would seem a non-trivial matter to understand how academics do conceptualise assessment. Their ideas can determine what assessment is and how it is conducted, which, in turn, can influence the quality of the learning in which students engage. It is against this background that the study reported here was carried out. The study is an exploratory analysis of the extent to which academics' perceptions of assessment are consistent with the construct of authenticity. The heuristic purpose of the study means that there is no claim for the findings to have generalisability to other contexts.

Introduction

Assessment tasks communicate to students the kind of intellectual work which is valued and thereby influence the way students behave as learners. Since assessment tasks drive the quality of student learning such that it is as high (or as low) as the cognitive demand level of the assessment tasks themselves (Crooks, 1988; Ramsden, 1997; Gibbs, 1999), then it would seem to be important that tasks are selected or designed so that the intended learning does indeed take place. The realisation that assessment is significant has been dawning in concert with reforms in assessment which shift the emphasis from an exclusive evaluation of declarative knowledge to include the assessment of procedural, strategic and conditional knowledge and understanding (Wiggins, 1989; Berlak, 1992; Nichols, 1994; Taylor, 1994). Essentially assessment tasks should reflect the ways in which knowledge and skills are used in real world contexts (Berlak, 1992; CSUP, 1992; Dearing, 1997). This means that assessment is concerned with the level to which knowledge is embedded in deep understanding, and can be demonstrated authentically (Newmann & Archbald, 1992; McDowell, 1998).

According to Wiggins (1989, 1993), authentic assessment tests our intellectual ability, or our ability to understand. It is only when we can use our knowledge 'wisely, fluently, flexibly and aptly in particular and diverse contexts' (Wiggins, 1993, p. 200) that we can be said to understand. The reproduction of cued knowledge, the application of algorithms and the performance of drills and exercises do not of themselves reflect understanding: the quality that Wiggins argues is at the core of authentic assessment. For Wiggins, any number of assessment tasks could be authentic but the central criterion that any task must meet for it to be classed as

Table 1. Standards by which intellectual achievement is judged authentic

-
1. *Analysis*—The response reflects higher order thinking with content by organising, synthesising, interpreting, evaluating and hypothesising to produce comparisons, contrasts, arguments, application of information to new contexts, and consideration of different ideas or points of view.
 2. *Disciplinary concepts*—The response reflects an understanding of ideas, concepts, theories and principles that are central to the academic or professional disciplines into which the student is being inducted.
 3. *Elaborated written communication*—The response explains understandings and conclusions. It is clear, coherent and provides richness in details, qualifications and argument.
-

Source: Adapted from Newmann, Marks & Gamoran (1996) and Newmann (1997).

authentic is its fidelity to the real world ways in which knowledge is used in the discipline or field of study into which the student is being inducted. The centrality of the 'real world' dimension in Wiggins' (1989, 1993) conceptualisations of authentic assessment has resulted in authentic assessment being otherwise referred to as performance assessment (Baker *et al.*, 1993; Baker & O'Neil, 1994; Torrance, 1995; Biggs, 1999) and indeed Wiggins himself does use the term 'performance'. However, Wiggins (1998) is explicit that authentic assessment and performance assessment are not synonymous because, he claims, performance assessment tasks need not 'replicate the conditions in which adults are challenged in context' (p. 21). A number of corollaries follow from the significance that Wiggins (1989, 1993) places on the congruence between an assessment task and reality, and these he discusses in some detail. In the interests of economy of space these will not be elaborated on here. In summary, however, it is clear that authentic assessment must focus on assessing intellectual achievement and while this may be logistically problematic, Wiggins (1993, 1998) argues that the validity of authentic assessment is not to be compromised by potentially competing priorities such as the relative ease of assessing lower-order, decontextualised skills, the practical problems of large-scale assessment and psychometric interpretations of students' responses. This is not to suggest that Wiggins discounts the important issues of validity, reliability and practicability. Rather, he argues that ways need to be found to resolve current difficulties in order that authentic assessment can achieve its purpose of enabling and improving achievement. Newmann *et al.* (1996) and Newmann (1997) build on Wiggins' ideas—though the work of Archbald and Newmann (1992) and Newmann and Archbald (1992) on defining authenticity in the context of achievement predates the elaborated writings of Wiggins—to adumbrate three standards by which the assessment of intellectual achievement can be judged authentic (Table 1).

The multifaceted nature of authentic assessment, as characterised above, is entirely consistent with contemporary theories of learning and knowing in which the emphasis is on how knowledge is represented, organised and processed in the mind. These constructivist perspectives on learning imply that assessment needs to encompass the more complex aspects of achievement, to recognise (from the wealth of literature on expert-novice differences) what successful learning looks like and to

acknowledge the powerful role of metacognition in effective thinking and competent performance. Both Wiggins' (1989, 1993, 1998) conceptualisations of authentic assessment and the criterial elaborations of Newmann and his colleagues suggest that authentic assessment could influence student learning in ways which more readily promote deep learning and critical thinking. However, as Cumming and Maxwell (1999) point out, while the term authentic assessment has gained widespread use in education, the concept is interpreted in different ways. It was against this background that the study reported here was conducted. The question, which drove the study, was: 'To what extent are academics' perceptions of assessment consistent with authenticity?'

Method

Depth interviews were carried out with a representative group (heads of department, module leaders/course directors and probationary lecturers/part-time tutors) of 12 academics. They were from the education faculty of one Scottish university, participated by invitation, had a range of disciplinary backgrounds, such as history, psychology and art, and came from different fields of practice, such as mainstream and special education, sport and speech therapy. The aim of the interview was to explore respondents' conceptions of desirable assessment. In other words, interviewees were being asked not what their assessment practices were but what they thought their assessment practices should be. The interviewer, a research student experienced in interviewing as a means of gathering data, was provided with a schedule of open questions plus prompts and probes to use if initial questions yielded underdeveloped responses. The questions had been identified as worthy of exploration in an earlier actuarial survey of extant assessment practices (Maclellan, 2001). Interviews were tape-recorded and transcribed.

The interview centred on two interrelated issues:

1. desirable assessment tasks to use with our students, with questions on:
 - the purpose of assessment for undergraduate students within the faculty
 - who assesses
 - what type of learning to assess and what modes to use
2. desirable scoring methods to determine the quality of student work, with questions on:
 - the use of criteria
 - establishing and maintaining marker consistency
 - the nature, extent and circumstances of feedback.

Additionally there was some exploration of the strengths, weaknesses and difficulties of each since it is the nature of the task together with how the student's performance on the task is judged that jointly indicate what learning is deemed important (Boud, 1990; Newmann & Archbald, 1992). Transcribed interviews were read for statements that reflected the standards, listed in Table 1. While Newmann and his colleagues have generated different versions of their criteria for authentic assessment, the list of three global criteria of Analysis, Disciplinary Concepts and

Elaborated Written Communication (Newmann, 1997) were deemed to offer a reliable coding framework with minimal risk of overlapping categories (Anderson, 1997). Since there was to be no attempt to quantify the content of participants' responses, impressionistic judgements were made against each criterion. Where statements clearly reflected any standard the category 'explicit' was endorsed (Table 2). Where it seemed reasonable to infer that a particular standard was intended, the category 'implicit' was endorsed. Where there was no evidence of a particular standard being suggested, the category 'not detected' was endorsed. Examples of each category can be seen in Table 3. To manage the coding in as straightforward a way as possible, the judgement made in respect of each standard reflected the 'highest' level of explicitness evidenced, regardless of the quantity of evidence available. Thus a transcript that reflected both explicit and implicit references to Analysis would be categorised 'explicit'. For each transcript three judgements were made.

Results

The interviews were to elicit views on what might be considered desirable assessment tasks and scoring methods, since inferences about what learning is deemed important derive from the nature of the task together with how the student's performance on the task is judged. Responses to each of the main questions are summarised below. Italicised script indicates verbatim quote from transcript.

- Purpose of assessment

Most of the responses included multiple purposes for assessment, with only one interviewee determining that assessment was exclusively about grading. These included the motivation of students (*to aid learning*), the diagnoses of students' strengths and weaknesses and the provision of evidence with which to engage in course evaluation. Dominantly, however, the purpose referred to in every response was that of student achievement. This was, in a few cases, represented as grading or ranking (*to see whether students have achieved acceptable standards; benchmarking student achievement; to classify students*) but most interviewees elaborated on achievement (*to let students demonstrate what they can do; to see that the student is actually learning; to check that the skills and attitudes for students' particular professional choice are developing*) to characterise it along a range of dimensions. While gauging achievement was regarded as an important purpose, respondents said that other purposes had been, historically, viewed as less important. Thus it was now considered appropriate that *diagnostic information be used more effectively to support student learning, and that more time be made available for tutors to develop and refine modules in the light of student need.*

- Who assesses

The common desire in all of the responses was that while faculty tutors should continue to be the major players in the assessment game, otherwise *the institution is criticised for not dealing with its professional responsibilities, students should have some say in the assessment of their work so that the idea of assessment does not remain foreign and*

abstract to them. While self assessment and peer assessment were viewed as *strengths* in an assessment system, self assessment was seen as *incredibly difficult for students* and students were viewed as *not having the skills to assess each other—they all give them grade As.* On balance, the respondents concluded that in the interests of public accountability, self and peer assessment was to be lauded in formative, but not in summative, assessment.

- What type of learning to assess and what modes to use

The assessment of content knowledge was expressed as *useful, important and/or necessary* by all of the respondents. Additionally, the respondents expressed the desire that students be able to *fuse the knowledge base with practice* in order to deal appropriately with professional issues. The assessment of higher order skills of *analysis, justification and evaluation* was also mentioned by some to be important. A few respondents mentioned the importance of assessing interpersonal skills since these mediated *competent professional functioning*. All of the respondents acknowledged that students preparing for professional practice should be ‘reflective practitioners’ who could *use their critical thinking faculties* but suggested that *we are a long way from knowing what is meant by critical thinking* and that *current attempts to assess reflection are really fairly crude*. There was also some desire that current practices of privileging knowledge recall be attenuated—*Perhaps too much of our assessment is just about how much they know when there could be more self evaluation and self critical analysis*.

Consistent with the view that the learning was complex and multi-faceted, responses to the question of how to assess expressed the need for variety in assessment tasks with greater attention to be paid to oral and practical modes. Justification for variety appeared to be on two grounds. One seemed based on a compensatory model of education—*we use oral, written and practical modes so that those who have skills in only one mode are not disadvantaged*. The other type of justification was a concern for the validity of the assessment task—*if the aim of the module is that the student learns to swim, the student demonstrates this by a practical swimming exercise, not a written test!* The importance placed on oral assessment (such as presentations, participation in discussion groups), practical assessment (such as fieldwork experience, case studies, the creation of artefacts/materials) and on mixed oral-practical assessment (such as reporting on practical tasks, poster sessions, group projects) was attributed to the significance of these modes of operation in professional life. here

All but one of the respondents considered written assessment to be desirable with declared preferences for the traditional academic essay and/or the reflective practice assignment (Brown *et al.*, 1997). The written mode was seen as desirable *because the student has to achieve great clarity to get his message across in writing; defining terms and explaining yourself are important and are parts of many written tasks; the written mode lets you be sure that the student actually understands about the area; there are times when you do need students to prove what they know and are able to construct and present some kind of case in written form*.

Table 2. Reference to standards

Case	Analysis	Disciplinary Concepts	Elaborated Written Communication
1	Not detected	Explicit	Not detected
2	Explicit	Explicit	Not detected
3	Not detected	Explicit	Not detected
4	Not detected	Implicit	Not detected
5	Not detected	Not detected	Not detected
6	Not detected	Explicit	Explicit
7	Not detected	Explicit	Not detected
8	Implicit	Explicit	Not detected
9	Explicit	Explicit	Not detected
10	Implicit	Explicit	Not detected
11	Implicit	Explicit	Implicit
12	Not detected	Explicit	Not detected

- The use of criteria

All respondents indicated that explicit assessment criteria should be available to students and tutors at the point when the assessment task was issued. Most suggested that detailed module-specific criteria were useful and even necessary because *they are effective ways of making sure that the standards are being met and they enable the marking process to be carried out more quickly*. However, most of respondents advocating the use of specific criteria recognised that broad and general criteria were also needed to capture *the important competencies and what is important in understanding and in thinking*. Both broad and specific criteria were seen as the most useful and logical place from to which to begin feedback to students. A couple of respondents pointed out that in detailing specific criteria, *you simply get overlapping categories and you generate too many ideas for the tutor to keep in her head when she is marking*. A few of the respondents, however, argued that criteria themselves are value laden (*the criteria you choose reflect your position about what is valuable to learn*) so cannot cover every eventuality and are open to interpretation.

- Establishing and maintaining marker consistency

Achieving marker consistency was seen as desirable and while there were extant technical requirements to have fails and marginal cases marked by second assessors as a matter of routine, the respondents expressed a desire for *shared understanding of the learning outcomes and objectives deemed important*. The practice of having meetings to mark and discuss sample assignments was considered important. Additionally it was suggested that *a much simpler marking system (such as pass/fail/merit); standardisation in the marking and grading; the provision of clear marking rubrics; and greater use of double or treble marking* would result in greater consistency.

- The nature, extent and circumstances of feedback

Feedback was considered valuable in supporting students' learning. *Some form of*

Table 3. Examples of standards of authenticity extracted from interview transcripts

Explicit analysis*Example 1*

Many of our modules require simple regurgitation of facts but I want students to go beyond what they are given to analyse ideas so that they can justify their practice. I want assessment which properly taps the understanding and application of ideas. We should be assessing the student's own well thought out ideas which are formed through the analysis and evaluation of their reading.

Example 2

In using a problem solving framework the student has to analyse the problem, create possible solutions, home in on a good solution or even recognise a candidate solution. We tend not to ask students to solve problems or make decisions or create resources. Rather we tend almost to view such tasks as being of a lesser nature than a traditional writing task

Implicit analysis*Example 1*

You sometimes think you understand until you try it out in practice and realise that you don't understand. This means that 'theory' and 'practice' are completely unrelated rather than forming a coherent picture in the mind.

Example 2

Within the personal tutorials that we run you can see the student who has the brilliant ideas, who is reflective, analytical and able to think things through. But none of that is actually assessed.

Explicit disciplinary concepts*Example 1*

One of the things we should be assessing is the student's grasp of knowledge and the ability to process that knowledge. If students had a substantial background and content knowledge in Maths or Science or whatever, then they'd be better able to cope with what they're being asked to teach in schools so I think it's important that we do check on subject knowledge.

Example 2

Students should be reading more original research literature rather than typical student texts that present theoretical ideas as acceptable and non-problematic. The research literature communicates better the discourse associated with the field. It highlights that there are no easy answers and so we each have to reach our own personal understanding. Using the research literature allows us to engage the students in debate rather than feed them with distilled wisdom.

Implicit disciplinary concepts*Example 1*

I'm very keen on peer-critiquing. By questioning others' assumptions and beliefs, you learn also to question your own. Students can engage in this if you set things up to let them do so. The process of peer review that goes on in the real academic world is one that I want to induct students into.

Example 2

I would like to see undergraduates doing more thinking themselves. I'd like to move towards assessment that uses the student's brain rather than the tutor's brain.

Explicit elaborated written communication*Example 1*

The assessment task should not be a test of recall but a test of understanding and the ability to interpret. It is necessary that the students write but in that writing it is necessary that they explain their thinking.

 Table 3. Examples of standards of authenticity extracted from interview transcripts—*continued*

Example 2

In the science students it is the lack of logical writing that I find most depressing. Writing is very often at the standard of a good fifth year pupil but not for someone who has been at university for years.

Implicit elaborated written communication*Example 1*

Students' writing is very important because the student has to achieve great clarity to get his/her message across in writing. So in a sense the writing promotes the skill of being clear.

Example 2

The written account may be a less good representation of the students' understanding than their oral account. I think this comes about because the students' conception of what goes into the writing is different from the academics'. The student may think that there are some things which are not appropriate to write down.

written feedback was given as a matter of course on every completed assessment task and was judged by tutors to be supportive, helpful and critical. Respondents viewed feedback as a trigger to engage students in discussion on how to improve as well as providing the opportunity for students to reveal their misunderstandings and thereby inform tutors about their own teaching. While tutors recognised that it was particularly weak students who made use of face-to-face discussions, it was considered desirable that all students should initiate formative discussion with their tutors, even although this was seen as very time-consuming. The plea expressed by some of the respondents was that students needed to view assessment as a mechanism for charting their development rather than for checking that they have learned or are doing the work.

In summary, desirable assessment takes account of diagnostic and evaluative purposes, rather than focusing exclusively on student achievement important though this is. Students themselves can benefit from greater involvement in the assessment process. The learning to be assessed is complex and so must be captured using a variety of modes, although the written mode is still important. Making judgements as to the adequacy of the students' performance on assessment tasks is complex. Opinion was divided on the extent to which criteria enables the judgement. More planned 'marker training' might improve judgements.

Discussion

The purpose of this study was to gauge the extent to which academics' perceptions of assessment were consistent with the construct of authenticity. Casual reflection on the interview results suggests that respondents' views on desirable assessment were consistent with Wiggins' expositions. They recognised the importance of using assessment to enable rather than just measure learning; they distinguished between trivial and non-trivial learning. They recognised the importance of context and the need to assess through valid tasks. Finally they seemed to recognise the frailty of

human judgement and the need to cope with this through the use of criteria and the development of shared understanding. In all of their deliberations the respondents communicated their awareness of the need for assessment to be in the 'real world' as the following extracts suggest:

The student has to demonstrate that he/she can do what the qualified professional actually does. So, under, supervision, the student has to assess the client's needs and make plans for subsequent professional action or intervention. Moreover the student has to carry out this task within an appropriate time-span and in the process maintain the client's co-operation and participation.

Students should be engaged in self-assessment but obviously this assessment should be subject to quality control. One quality control mechanism is the peer reviewing process that goes on in journals. I think students should be introduced to some version of peer review.

When the external examiner comes to us and says 'the industry is looking for x' we have to examine whether the assessment demands that we make of students are appropriate. It may be that the industry wants students who are more proficient in the studio or in the practice rooms or who can produce quality artefacts more speedily but here in the Faculty we also have to consider whether or not the task makes academic demands so that the students actually reflect on the value of their creative/artistic product.

But critical reflection on the interview responses suggests that initial impressions need to be qualified. According to Wiggins and to Newmann *et al.*, the essence of authenticity is the intellectual achievement of using knowledge to address real world tasks and problems. Authentic assessment is thus trying to capture what Bereiter (2002) calls conceptual artefacts, by which he means the new conceptualisations that students develop in order to better understand 'real world phenomena and the texts that refer to them' (p. 294). Authentic assessment is thus trying to make a judgement about what is essentially cognitive activity. Because cognitive activity cannot be directly observed, assessing its incidence is a process of making inferences from observations and this, as the literature attests, is problematic. Notwithstanding the procedural and logistical constraints of authentic assessment, it is the essence of what authenticity *is* that is the focus of this study, and to make an evaluation of this, the standards adumbrated by Newmann seem appropriate criteria.

As can be seen from Table 2, all but one of the transcripts made reference to the need for content knowledge and/or the need to engage appropriately in the disciplinary process. The inclusion of disciplinary concepts as elements to assess was heartening since it is now well recognised that a developed knowledge base together with the methodological knowledge to investigate the domain is a mark of the expert rather than the novice (Glaser, 1984). Because experts are not people who have all the answers but rather are trying to solve new problems, they need to create new knowledge to solve the problem. In order to do this they need a well-developed knowledge base and knowledge of methods of enquiry, research or communication with which to build on the extant knowledge. It is the product of the interaction between extant knowledge and the methodological means to improve/refine this knowledge that is the essence of authentic achievement (Archbald & Newmann,

1992; Newmann & Archbald, 1992). That assessment should reflect knowledge and thinking as it is needed in the professional field is exemplified in Table 3.

However, the interaction between extant knowledge and the methodological means to refine such knowledge cannot take place without elaborated written communication, a criterion to which only two respondents made explicit reference. Yes, there were many expressions which asserted that written assessment was desirable, but respondents often elaborated such assertion in ways that belied their initial expression. For example, while conceding that the *written mode of assessment allows students to reveal their understanding*, several of the same respondents also declared that *talking with the students is more likely to give you a better picture of whether they understand* and that *the oral mode has the advantage of allowing students to reveal their thinking*. Most respondents viewed the written mode as the means *for students to prove what they know and write it down in an academic way*. Indeed *if the assessment is on material that is factual and structured then we should require a written piece of work*. Only two of the respondents conceptualised writing as a tool for thinking and making meaning (Kellog, 1994). One explanation for this is that many of the respondents may have a knowledge-telling view of writing (Bereiter & Scardamalia, 1987) as will be outlined below.

There is considerable encouragement for higher education to increase diversity in assessment methods (Yorke, 1998; Biggs, 1999; Race, 1999). The rationale for this appears to be that since methods with which the student is uncomfortable may handicap the student and thereby compromise the validity of the assessment, it is appropriate to select methods through which the students can comfortably demonstrate their learning. Given that writing tasks are problematic for many students in higher education (Lea & Stierer, 2000), it would not be unreasonable if tutors (with kindly but misplaced intentions) eliminated or at least attenuated the writing demands they make of students. Indeed there is some suggestion of this by the interviewees. For example, one asserted that

It can be difficult to determine how much a student understands when reading the typical written assignment because it favours students with an arts background whilst students with a science background may be quite unpractised in essay writing.

Another explained that

We encourage students who may not have standard entry to come into the department and want all of them to work to their full potential. We find that by having a variety of forms of assessment, students have the opportunity to work through the course in an appropriate manner. Students who find writing difficult can therefore find a different way of expressing themselves.

However, whilst there are very good reasons for enabling students in the assessment process, it is important not to discount the value of writing in enabling learning. Bereiter and Scardamalia (1987) make a clear distinction between writing which records the thoughts that are already in one's mind (characterised as the knowledge-telling model) and writing which is shaped to achieve specific purposes and which reorganises one's knowledge in the process (characterised as the knowledge-transforming model). The knowledge-transforming model subsumes (but is not merely

an elaboration of) the knowledge-telling model, and notwithstanding the everyday situations for which the knowledge-telling model is perfectly adequate, it is the knowledge-transforming model which is the more powerful because it enables us to rework our thoughts. In using the knowledge-transforming model we can consider whether the text we have written says what we want it to say and indeed whether we believe what the text says. If, however, tutors are unaware of, or unconvinced by, the idea of knowledge-transformation (which is, after all, only a psychological construct invoked to explain what is believed to be involved in the composing process) but operate, even implicitly, with the model that writing is only about knowledge telling (in which the essential elements are the retrieval of content from memory and knowledge of the appropriate literary form in which to write), then they are unlikely to be persuaded by the need to clarify and extend one's mental representations of ideas.

While elaborated writing is the mechanism through which disciplinary content knowledge and process are depicted, the cognitive effort involved is the higher order thinking that Newmann characterises as analysis. Table 3 shows examples of what explicit and implicit references to analysis looked like and Table 2 shows that the standard was recognised by five of the respondents. This suggests that for most of the respondents the construction of meaning was not a mark of learning. Such a finding is alarming because the current emphasis on student-centred learning (Bowden & Marton, 1998; Biggs, 1999; Gibbs, 1999; Prosser & Trigwell, 1999) presupposes models of teaching and assessment that are consistent with constructivism. Constructivism considers knowledge to be created as persons adapt to, and make sense of, their experiential worlds (von Glaserfeld, 1987, 1996). Given that our own actions and our reflections on our actions are 'constructed' as distinct from 'given', it would seem to be important that our perceptions of similarities and differences in relation to experiences, our evaluations of the experience and our abstractions from specific experiences into superordinate generalisations were sincerely considered as evidence of our learning. One conclusion to derive from the respondents' disregard for the need to construct knowledge is that the many current allusions to the importance of 'student-centred' learning are based on rhetoric rather than theoretical insight.

Conclusion

The purpose of the study reported here was to explore how authentic tutors' perceptions of desirable assessment were. To the extent that the interview responses only partially met the three criteria for authentic assessment, the construct of authenticity is not being fully realised. While it may be unreasonable to expect all assessment tasks to meet all three criteria all of the time (Newmann *et al.*, 1996), the assessment of discrete knowledge and skills which may well be foundational to authentic achievement does not of itself constitute authentic assessment. Authentic assessment is trying to judge what genuine knowledge building there has been (in the context of a real world problem or concern) and the evidence for this comes from what Bereiter (2002) calls progressive discourse. While the respondents'

discourse included reliance on disciplinary content knowledge and process, it was much less concerned with either the conceptual tools of analysis or the mechanism of extended writing. In other words while the respondents used language which could, superficially, be taken to refer to authenticity, their own understandings of the construct were not comprehensive. While mastery of disciplinary concepts is necessary in authentic assessment, it is not sufficient. The interdependence of disciplinary concepts, elaborated writing and analysis has to be acknowledged and operationalised for assessment to be tapping the real world understanding or intellectual achievement which Wiggins argues is so important. It would seem, from the findings of this exploratory study that considerable curricular development may be needed if academics are to realise authentic assessment as intended by Wiggins.

References

- Anderson, J. (1997). Content and text analysis. In J. Keeves (Ed.), *Educational research, methodology and measurement: An international handbook* (pp. 340–344). Oxford: Elsevier Science Ltd.
- Archbald, D., & Newmann, F. (1992). Approaches to assessing academic achievement. In H. Berlak, F. Newmann, E. Adams, D. Archbald, T. Burgess, J. Raven, & T. Romberg (Eds.), *Towards a new science of educational testing and assessment* (pp. 139–180). New York: State University of New York Press.
- Baker, E., O’Neil, H., & Linn, R. (1993). Policy and validity prospects for performance-based assessment. *American Psychologist*, 48, 1210–1218.
- Baker, E., & O’Neil, H. (1994). Performance assessment and equity. *Assessment in Education*, 1, 11–26.
- Bereiter, C. (2002). *Education and mind in the knowledge age*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Bereiter, C., & Scardamalia, M. (1987). *The psychology of written composition*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Berlak, H. (1992). The need for a new science of assessment. In H. Berlak, F. Newmann, E. Adams, D. Archbald, T. Burgess, J. Raven, & T. Romberg (Eds.), *Towards a new science of educational testing and assessment* (pp. 1–21). New York: State University of New York Press.
- Biggs, J. (1999). *Teaching for quality learning at university*. Buckingham, UK: The Society for Research into Higher Education & The Open University Press.
- Boud, D. (1990). Assessment and the promotion of academic values. *Studies in Higher Education*, 15, 101–111.
- Bowden, J., & Marton, F. (1998). *The university of learning*. London: Kogan Page.
- Brown, G., with Bull, J. & Pendlebury, M. (1997). *Assessing student learning in higher education*. London: Routledge.
- Committee of Scottish University Principals. (1992). *Teaching and learning in an expanding higher education system* (The MacFarlane Report). Edinburgh: CSUP.
- Crooks, T. (1988). The impact of classroom evaluation practices on students. *Review of Educational Research*, 58, 438–481.
- Cumming, J., & Maxwell, G. (1999). Contextualising authentic assessment. *Assessment in Education*, 6, 177–194.
- Dearing, R. (1997). *National committee of inquiry into higher education* (Dearing Report), Higher Education in the Learning Society, Report of the National Committee, Norwich: HMSO.
- Eraut, M. (1994). *Developing professional knowledge and competence*. London: The Falmer Press.

- Gibbs, G. (1999). Using assessment strategically to change the way students learn. In S. Brown & A. Glasner (Eds.), *Assessment matters in higher education*. Buckingham, UK: The Society for Research into Higher Education & The Open University Press.
- Glaser, R. (1984). Education and thinking: The role of knowledge, *American Psychologist*, 39, 93–104.
- Hativa, N. (2000). Becoming a better teacher: A case of changing the pedagogical knowledge and beliefs of law professors, *Instructional Science*, 28, 491–523.
- Kellog, R. (1994). *The psychology of writing*. Oxford: Oxford University Press.
- Kember, D. (1997). A reconceptualisation of the research into university academics' conceptions of teaching. *Learning and Instruction*, 7, 255–275.
- Kember, D. (1998). Teaching beliefs and their impact on students' approach to learning. In B. Dart & G. Boulton-Lewis (Eds.), *Teaching and learning in higher education*. Melbourne: The Australian Council for Educational Research Ltd.
- Lea, M. & Stierer, B. (2000). *Student writing in higher education*. Buckingham, UK: The Society for Research into Higher Education & The Open University Press.
- Maclellan, E. (2001). Assessment for learning: The differing perceptions of tutors and students. *Assessment & Evaluation in Higher Education*, 26, 307–318.
- Martin, E., Prosser, M., Trigwell, K., Ramsden, P., & Benjamin, J. (2000). What university teachers teach and how they teach it. *Instructional Science*, 28, 387–412.
- McDowell, L. (1998). Editorial, *Assessment and Evaluation in Higher Education*, 23(4), 335–338.
- Newmann, F. (1997). Authentic assessment in social studies. In G. Phye (Ed.), *Handbook of classroom assessment* (pp. 359–380). London: Academic Press.
- Newmann, F. & Archbald, D. (1992). The nature of authentic academic achievement. In H. Berlak, F. Newmann, E. Adams, D. Archbald, T. Burgess, J. Raven, & T. Romberg (Eds) *Towards a new science of educational testing and assessment* (pp. 71–83). New York: State University of New York Press.
- Newmann, F., Marks, H., & Gamoran, A. (1996). Authentic pedagogy and student performance. *American Journal of Education*, 104, 280–312.
- Nichols, P. (1994). A framework for developing cognitively diagnostic assessments. *Review of Educational Research*, 64, 575–603.
- Pajares, M. (1992). Teachers' beliefs and educational research: cleaning up a messy construct. *Review of Educational Research*, 62, 307–332.
- Prosser, M., & Trigwell, K. (1999). *Understanding learning and teaching*. Buckingham, UK: The Society for Research into Higher Education & The Open University Press.
- Race, P. (1999). Why assess innovatively? In: S. Brown & A. Glasner (Eds.), *Assessment matters in higher education* (pp. 57–70). Buckingham, UK: The Society for Research into Higher Education & The Open University Press.
- Ramsden, P. (1997). The context of learning in academic departments. In F. Marton, D. Hounsell, & N. Entwistle (Eds.), *The experience of learning* (pp. 198–216). Edinburgh: Scottish Academic Press.
- Roth, R. (1999). *The role of the university in the preparation of teachers*. London: Falmer Press.
- Taylor, C. (1994). Assessment for measurement or standards: the peril and the promise of large scale assessment reform. *American Educational Research Journal*, 31, 231–262.
- Torrance, H. (1995). Introduction, in: H. Torrance (Ed.), *Evaluating authentic assessment* (pp. 1–8). Buckingham, UK: The Open University Press.
- von Glaserfeld, E. (1987). Learning as a constructive activity. In C. Janvier (Ed.), *Problems of representation in the teaching and learning of mathematics*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- von Glaserfeld, E. (1996). Aspects of constructivism. In: C. Twomey Fosnot (Ed.), *Constructivism*. New York: Teachers College Press.
- Weil, S., & McGill, I. (1989). *Making sense of experiential learning*. Buckingham, UK: The Society for Research into Higher Education & The Open University Press.
- Wiggins, G. (1989). Teaching to the (authentic) test. *Educational Leadership*, 46, 41–47.

- Wiggins, G. (1993). Assessment: authenticity, context, and validity. *Phi Delta Kappan*, 75, 200–214.
- Wiggins, G. (1998). An exchange of views on semantics, psychometrics and assessment reform: a close look at authentic assessments. *Educational Researcher*, 55, 19–22.
- Yorke, M. (1998). Assessing capability. In J. Stephenson & M. Yorke (Eds.), *Capability and quality in higher education* (pp. 174–191). London: Kogan Page.

