The Enhancement of Teaching and Learning in Higher Education through Accreditation, Portfolio Assessment, Educational Development Projects and Evaluation using Cyclical Models of Learning

Presented for the award of PhD by Publication
at The University of Glamorgan
November 2004

David Baume
64 Princess House
144 Princess Street
MANCHESTER
M1 7EP

Phone 0161 273 6454
Mobile 07747 045 931

E-mail ADBaume@aol.com
Acknowledgements

I thank:

All my friends and colleagues in the staff and educational development community, in the UK and beyond, over the last fifteen years or so, for their generosity and optimism and talent and effort in bringing staff and educational development from a cottage industry to, if not exactly mainstream, then certainly a fast-flowing, oxygen-rich and sparkling tributary of the mainstream of higher education. They are the nicest and most collaborative group of professionals I have ever worked with.

Those in higher education who, whilst not developers, became enthusiasts for our work; at all levels, from Vice-Chancellor up to teaching assistant and student.

My co-workers and co-authors in the work included in this submission - Carole Baume, Alan Jenkins and Mantz Yorke - for many hours of challenging, productive and enjoyable thinking, talking, writing and then editing and swapping of drafts.

Colleagues and friends for discussion variously of ideas, outlines and drafts of this Overview – Roni Bamber, Phil Candy, John Cowan, Robert Edwards, Peter Knight, Jo Tait, Rosamund Woodhouse and, again, Mantz Yorke.

My mother Anne, for insisting that I take a PhD so that she could at last come to one of my degree ceremonies.

My partner Carole, again, for support and tolerance and colleagueship beyond description or measure.

The offspring (no longer children): Beth for repeated doses of “Of course you can do it”; Kit for the Springsteen quote, and also for pointing out that my evolving series of questions constitutes its own cycle of development; and Rich for sustained interest and confidence in my completion and success.
Christine Sneed for rigorous and supportive proofreading.

My supervisors, Professor Danny Saunders and Professor David Turner, for just the right combination of guidance, feedback, challenge, confidence and patience.
Contents

1 Introduction: Historical background

2 Overview

2.1 Theoretical context

2.2 Improving learning and teaching through accreditation

2.3 Improving learning and teaching through portfolio assessment

2.4 Improving learning and teaching through educational development projects

2.5 Improving learning and teaching through evaluation

2.6 Kolb and intentional learning

2.7 Originality and impact

2.8 Purposive learning from experience

3 References

4 The evidence

Tabulation of themes against publications

Appendices

1 Citations and reviews

2 Co-worker statements

3 Overview of work on accreditation
1 Introduction: historical background

This introduction summarises the development of my work and thinking in the four areas which form the focus of this PhD – Accreditation of Teachers in Higher Education, the Assessment of Portfolios, Educational Development Projects, and Evaluation.

All of the work reviewed here is rooted in a concern for learning and its improvement. My work and research has progressed through six layers or levels of learning during my career in higher education. These layers and levels, and the associated questions, are listed in Table 1.1 and explored further below:

<table>
<thead>
<tr>
<th>Layer or level</th>
<th>Associated Question</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>“Under what conditions do students learn?”</td>
</tr>
<tr>
<td>2</td>
<td>“What actions by teachers support or impede student learning?”</td>
</tr>
<tr>
<td>3</td>
<td>“Why do lecturers teach as they do, given what is known and published about the conditions for effective learning in Higher Education?”</td>
</tr>
<tr>
<td>4</td>
<td>“What actions by developers (by me) support or impede staff learning about teaching?”</td>
</tr>
<tr>
<td>5</td>
<td>“How can we develop policies, strategies, environments and systems within which teachers not just can teach well but do teach well, and students can and do learn well.”</td>
</tr>
<tr>
<td>6</td>
<td>“How can I help national higher education systems, agencies, organisations, institutions, professions and individuals to provide the conditions under which all of the above learnings will happen?”</td>
</tr>
</tbody>
</table>

Table 1.1 The layers, levels and progression of my work in higher education, and the associated questions

Moving on to the later questions did not mean stopping addressing the earlier questions – my current work on questions 3-6 is still powerfully informed by work on questions 1 and 2.

First, the learning of students. I started asking "How do I learn?" and "Under what conditions?" as an undergraduate in the 1960s, when I was
subject to teaching across the full range from execrable to brilliant. Execrable included the physics lecturer who each week read to us, without hesitation or interruption, a chapter from his textbook, which he had previously insisted we each buy. Brilliant included listening each week to Karl Popper as he expounded on some problem of philosophy and ways to address it. As an undergraduate I tentatively concluded that learning needed engagement with, even excitement about, what was being learned; clarity about the point of the learning, the question or problem being addressed; and learner activity, which meant at least some form of mental activity rather than the mere taking of dictation. These early answers, albeit obtained through introspection rather than research, have served me well.

Second, starting in 1970, I became interested in my own learning as a lecturer. The question “Under what conditions do students learn?” can be reframed from the point of view of a lecturer as “What actions by the lecturer support or impede student learning?” At first, my teaching covered much the same range of quality that I had experienced as a student. Execrable included lecturing on irrelevant economics to six justifiably bored HND engineering students.

Excellent included persuading fellow lecturers to allow an exceptionally able corps of first year mechanical engineering students to undertake a first-year project, designing, building and operating in the open sea a piloted submarine camera platform instead of doing the normal laboratory work, complementing staff’s research into underwater work systems. Exploring relationships between research and teaching, Healey (2005 – in preparation), reports that “... students are likely to gain most benefit from research, in terms of depth of learning and understanding, when they are also involved in research, for example, through various forms of active learning, such as inquiry-based learning”, a finding strongly supported by this experience. Healey also comments: “This presents challenges to university staff to reshape curricula and may lead to new ways for staff and students to work together...” The challenges presented to my academic
colleagues were indeed so great that this curricular experiment was never repeated, although it did generate new ways of working, in particular a closeness of collaboration among staff and students that we were able to replicate through smaller projects and practical diving fieldwork.

Research and publishing, about the learning and teaching of engineering as well as about research within engineering, were part of my lecturing career almost from the start - see for example Baume and Jones 1973, 1976; Hipwell and Baume 1976. Much of this educational research and development involved applying and (like a good Popperian) testing and refining earlier conclusions from asking question about my learning as a student. I additionally became impressed by the contribution that goals expressed as clear learning outcomes (then called objectives) made to learning, teaching and assessment (Baume and Jones 1974). Questioning became much more sophisticated through completing a part-time MA in Curriculum Development in Higher Education in 1978.

Research interviews for the MA (Baume 1977) led to the view that lecturers in engineering had two professions, engineering and lecturing, and further to the view that these two professions are of very different kinds, albeit potentially complementary. “Engineering is what engineering lecturers know about; lecturing is what they do with it”, as I rather crudely but nonetheless fairly put it (ibid p 56).

Some ten years after this work on lecturers in engineering, in the late 1980s, and after some 20 years of lecturing, I shifted to my third focus, the learning of other lecturers, starting to ask "Why do lecturers teach as they do, given what is known and published about the conditions for effective learning in Higher Education?" Increasingly intrigued by this question, I became a Staff and Educational Developer. Sometimes, those early days of work as a developer were very dark. For example, no one at all came to the first workshop that I organised with an outside speaker. That day, Chris Rust from the Oxford Centre for Staff Development and I ate a lot of sandwiches. Together we despaired at the way that lecturers can teach in an almost
wholly a-theoretical way that would be anathema within the practice of their discipline. This theme is revisited later.

Working as a developer led to my asking a very similar question about lecturers' learning to those asked earlier about students' learning - "Under what conditions do lecturers learn (about teaching)?"

Fourth, and very quickly after question three, and following what was by now becoming a clear, even predictable, progression, came the question "What actions by developers support or impede staff learning about teaching?" Throughout my work as a developer I have been concerned with this question. Some of my early answers as a student and then as a lecturer - about clear goals, about active engagement - turned out also to apply to work as a developer. However, the situation is greatly complicated: By the relations between academics' two professions (their subject and the teaching of their subject), which are much more complex that my 1977 analysis suggested: By institutional and professional recognition and reward structures: And by national factors such as the Teaching Quality Enhancement Fund, the emergent CETLs, and by successive Research Assessment Exercises, which greatly affect institutional and individual priorities as between teaching and research.

Fifth, realising the vital importance of context for all of my work and for all of these previous questions, and prompted by the realisation that institutions as well as individuals learn, I asked "How can we develop policies, strategies, environments and systems within which teachers not just can teach well but do teach well, and students can and do learn well?" This led me, as a developer and development manager, to contribute to university policy and strategy that valued and supported learning and teaching in locally appropriate ways. It also led me to help others to do the same.

Sixth, by scaling up the previous question, we get: "How can we help national higher education systems, agencies, organisations, institutions, professions and individuals to provide the conditions under which all of the
above learnings will happen?" I addressed this question even before I had explicitly asked it. I have continued to address it through work on accreditation from 1990 to the present, and will continue to do so on into the future (See Appendix 3).

Questions 3 to 6 have driven much of my work in academic development. I should explain what I mean by ‘academic development’. Academic development was described as clearly as possible in a final editorial for the International Journal for Academic Development (IJAD). (I co-founded IJAD in 1996 and co-edited it until 2002. It is the first, and still the only, journal in the field.): “Academic development is concerned with the improvement of the processes of higher education – educational development – and with enhancing the capabilities of those who directly support learning in higher education – staff or faculty development” (Baume 2002:2 – emphasis added).

This sequence of questions shows a steadily broadening focus, from personal to local to institutional to professional to national and in some respects to international. But the intent of each question is much the same – to understand in order to be able to act more effectively.

There has been a serious academic frustration in much of this work, only recently being resolved. Teaching and academic development in higher education have mostly been empirical affairs. Initially, this empirical approach reflected a comparative lack of published research into learning and teaching in higher education. But slowly, the research was undertaken. An excellent early review on learning from lecturing was Bligh (1971), recently updated (Bligh 2000). But the research was mostly ignored. Academics were often reluctant to, as they sometimes expressed it, learn another subject – higher education.

Developers found ourselves in a bind. Should we base our work on useful but unreferenced hints and tips, thus seeming to ignore the research underpinnings and perpetuating the idea that teaching and learning are
largely a-theoretical? (‘Seeming to’ because much of the hints and tips literature is in fact strongly, albeit implicitly, underpinned by theory and scholarship.) Or should we make our sources explicit, and risk being told that our ‘eduspeak’ was irrelevant to discipline X, Y or Z? Often we chose the former. (We were not in the same bind about academic development, where until the last few years there has been little or no research to use.)

The situation has changed over some 15 years, and is changing with increasing speed. There is a fast-growing literature about teaching and learning in higher education, both generic and more recently in a growing range of disciplines. The latter is exemplified by the RoutledgeFalmer ‘Effective Learning and Teaching in Higher Education’ series developed with the Institute for Learning and Teaching in Higher Education (ILTHE) / The Higher Education Academy (HEA). And this literature is increasingly being used, to support educational development projects and on accredited courses in teaching in higher education. For example in Baume (1998:1 – Publication 3) it felt appropriate to use some ten references in a guide to lecturers on building their teaching portfolio. But there is a very long journey to travel. Developers, too, have a fast-growing scholarly as well as practical base for our work (see for example Wisdom and Macdonald (2001); Edwards, Baume and Webb (2003); Kahn and Baume (2003); Eggins and Macdonald (2003), Baume and Kahn (2004)). The pleasing resolution of the frustration over the often secret scholarship of teaching and development is that teachers and developers increasingly feel able to come out of the closet, to be explicitly scholarly about their teaching and development work. This emergence of explicit scholarship is another theme in the work described in the Overview.

This work makes much reference to learning. The learning of an individual or organisation is considered here to be an increase in their or its capabilities to achieve their or its goals. This account of learning explicitly looks at learning through the uses to which the learning can be put, because that is the kind of learning in which I am most interested. (Why ‘capabilities to
achieve goals’ rather than the more important ‘achieving of goals’? Because obstacles beyond the reach and control of the learner can sometimes impede their actual attainment of their goals.) In the occasional absence of explicit goals, learning here means an increase in the number of things (of all kinds, including mental, physical, social, emotional) that can be done, and also in the quality – which may mean complexity, appropriateness, versatility – with which they can be done.

Driving my interest in and engagement with each of these questions, for most of my career, has been a commitment to applying an evolving understanding of learning to the improvement of learning.
2 Overview

2.1 Theoretical context

A PhD "is a significant step forward in the work done in a specific area which is disciplinary or interdisciplinary based." (Wisker 2001:23) That step forward may take the form of a purposive addition to and / or refinement of the current state of knowledge or understanding in a specific area. Brew (2001:7) is right to say that the "... creation of knowledge [has] become contested space". But, despite this, the academic enterprise continues. What are the capabilities of a PhD-holder? "Holders of doctorates will be able to conceptualise, design and implement projects for the generation of significant new knowledge and/or understanding. Holders of doctorates will have ... the ability to make informed judgements on complex issues in specialist fields, and [to innovate] in tackling and solving problems." (QAA 2001)

The work reviewed here describes learning, research, progress, and the creation, interpretation and use of significant (valid, accepted, productive) knowledge. Much of the research, the production of knowledge, reviewed here has been undertaken in what Gibbons et al. (1994) calls Mode 2. In Mode 2, knowledge is produced, in an informed, scholarly and reflective way, in and through action. Gibbons et al. contrast Mode 2 with the classic model of Mode 1 knowledge production, in which knowledge is first produced through research and then, as a separate act, applied to practice. Mode 2 knowledge production also sits very comfortably with the view of learning as an active process where the learner constructs their own models and understandings rather than somehow absorbing those that they are taught (e.g. Vygotsky 1986).

Below are considered various accounts of the relationships between learning, research and the production of knowledge, including the work of Nonaka (1994), Gibbons et al. (op. cit.) and Kolb (1984), Kolb et al. (1991)
with reference also to the work of Vygotsky (1986), Schön (1982 and 1987), Bruner (1977), and Ashwin and Trigwell (2004). The Overview combines and re-synthesises aspects of all this work, and shows cycles of goal-directed activity in which knowledge is produced, usually through action; made explicit; tested and refined; and then used. The knowledge production considered here is a cyclical process; more precisely a spiral process, each stage and then each cycle building on the previous. It involves both individual and collaborative work.

2.1.1 Nonaka

For Nonaka (1994), knowledge is produced, by individuals working in organisations, through a sequence – more usefully be seen as a spiral - of continuing conversations about their work and about their understanding thereof. In these conversations they first share their tacit knowledge (in cell 1, ‘socialisation’). They transform this shared tacit knowledge into explicit knowledge (in cell 2, ‘externalisation’). They share and test and resynthesise with colleagues this now explicit knowledge (in cell 3, ‘combination’). And they convert it back into more sophisticated tacit knowledge which informs their improved practice (in cell 4, ‘internalisation’). Two half-cycles of this, from different stages of my work, are shown in Table 1.2:

<table>
<thead>
<tr>
<th>Conversion of Tacit Knowledge</th>
<th>Explicit Knowledge</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tacit Knowledge</td>
<td>Tacit Knowledge</td>
</tr>
<tr>
<td>1 Socialization – we share our tacit knowledge (it remains tacit). <strong>As in</strong> my earliest work with colleagues developing new teaching methods. This socialisation started with us simply talking, mostly a-theoretically, about our teaching.</td>
<td>Explicit Knowledge</td>
</tr>
<tr>
<td>2 Externalization – we talk with others to help us and them to externalise, to make explicit, our tacit knowledge. <strong>As in</strong> writing our first publications and presentations about teaching and learning, describing our practice and the reasons for and effects of it (e.g. Hipwell and Baume 1976).</td>
<td>Explicit Knowledge</td>
</tr>
</tbody>
</table>


Table 1.2 – Nonaka’s account of the production of knowledge in organisations, applied to two parts of my own work

2.1.2 Gibbons et al., Ashwin and Trigwell

The work of Gibbons et al. is summarised at the start of 2.1. Mode 2 is exemplified in action research, where action and research, the production and use of knowledge, are inseparable. Mode 2 knowledge production also sits very comfortably with the view of learning as an active process, a process in which the learner constructs their own models and understandings of the world rather than somehow absorbing those that they are taught (e.g. Vygotsky 1986).

Education has mostly worked in Mode 1. Once Mode 1 knowledge has been produced, it is also taught, from the primary research publications or from secondary overview sources such as books (including textbooks), lecture notes or on-line resources; hopefully learned; and then hopefully applied to academic or professional practice. It may be argued that Mode 1 is as problematic, and often as inappropriate, in teaching as in research.

Ashwin and Trigwell (2004) usefully enrich Gibbons et al.’s account, distinguishing between three possible purposes for investigating (i.e. learning and generating knowledge about) professional practice. The first is to generate personal knowledge, to inform one’s own practice. The second
is to generate local knowledge, tested by and then informing the work of close colleagues. The third is to generate public knowledge, tested through refereeing and then used to inform the wider professional community. The latter process is normally called research.

These three stages may, but do not necessarily, form a developmental sequence. Ashwin and Trigwell account well for the developmental process described here and in the papers in my work in teaching and development, although their first two stages have mostly been elided, as most of my early teaching development was undertaken working closely with fellow teaching enthusiasts. But my research arises from my practice, and is undertaken in and about that practice. Its primary intention is to inform practice, mine and that of colleagues and more distant others. Lewin’s view that "There is nothing so practical as a good theory" (Lewin 1951:160) resonates very strongly. But I work and research in Mode 2, driven by the developer’s intention to make things better. A necessary part of this quest, of course, is to produce solid and applicable knowledge and understanding, that is, to do research.

2.1.3 Kolb, Schön and Cowan

Learning and the production of knowledge can usefully be seen as a four-stage cyclical – again more accurately a spiral – process; experience, reflection, and seeking new and more advanced understanding and knowledge, followed by applying and testing this (hopefully) advanced understanding and knowledge in future planning and action (Kolb 1984; Kolb et al. 1991 pp 59-60). Bruner (1960) is explicit about learning being a spiral process – indeed he described a form of curriculum, the spiral curriculum, in which learners are encouraged to revisit concepts at increasing levels of sophistication.
Concrete experience

Testing implications of concepts in new situations

Formation of abstract concepts and generalisations

Observation & reflection

Figure 2.1 Kolb’s account of learning from experience (Kolb, Rubin et al. 1991 p 59)

Concrete experience

Active Experimentation

Reflective Observation

Abstract Conceptualisation

Figure 2.2 Kolb’s account of learning from experience (Kolb 1984)

The revised and simplified version below is very useful:

Figure 2.3 A revised account of Kolb’s Learning Cycle to show learning from action intended to lead to learning

The revised cycle in Figure 2.3 describes action intended to lead to learning. It can be joined anywhere. The plan is to achieve something in the world, to learn, hopefully to do both. After doing comes reflection or review, identifying how effective the plan and the consequent action were, followed
by theorising, searching for explanation to inform planning for the (hopefully) more effective next plan and next action – and so on.

In Table 2.3 the two cycles are presented alongside each other, with comments:

<table>
<thead>
<tr>
<th>Kolb's original cycle</th>
<th>Revised version</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Do</td>
<td>Concrete experiences</td>
<td>Do</td>
</tr>
<tr>
<td>O</td>
<td>Observation and reflection</td>
<td>Review</td>
</tr>
<tr>
<td>O</td>
<td>Formation of abstract concepts and generalisations</td>
<td>Explain</td>
</tr>
<tr>
<td>O</td>
<td>Testing implications of concepts in new situations</td>
<td>Plan</td>
</tr>
</tbody>
</table>

*Table 2.3  A comparison of Kolb's and the revised account of learning from experience*

Table 2.4 applies the model to the experimental work on portfolio assessment described in Baume and Yorke (2004 – Publication 10):

<table>
<thead>
<tr>
<th>Project</th>
<th>Stage of revised cycle:</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Portfolio assessment experiment paper (Baume and Yorke 2004)</td>
<td>Plan</td>
<td>Do</td>
<td>Review</td>
</tr>
<tr>
<td>Plan</td>
<td>Do</td>
<td>Review</td>
<td>Explain</td>
</tr>
<tr>
<td>1 Review previous published paper on portfolio assessment (Baume and Yorke 2002), and realise that an important question remains unanswered – why did the assessors make the judgments that they made?</td>
<td>2 Develop possible explanations, but realise that further research is needed</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 2.4 – The revised learning cycle applied to an experiment on portfolio assessment

Note that the account in Table 2.4 starts at the review stage, and requires additional steps around the cycle to describe the work completely.

The formal, step-by-step process described here is characteristic of what Dreyfus and Dreyfus (2000) would characterise as novice, or perhaps at the most advanced beginner, behaviour. Is it appropriate to ascribe such a step-by-step, novice-like process to competent, proficient and expert behaviour? I suggest that it is, and for three reasons. These reasons are best illustrated through a personal account. First, I claim expertise in many areas of my work, and yet I still undertake these four steps systematically, albeit hopefully undertaking each step in an increasingly efficient and sophisticated way. Second, most of my work, like the work of many professionals, is being done for the first time – for the first time in these particular settings and with these particular goals or constraints (such as running a class or workshop), or sometimes for the first time ever (devising an accreditation framework for a particular profession). These three facts makes me a perpetual novice, although a novice inventor rather than a novice learner of the already known, living forever in my zone of proximal development (Vygotsky 1978), often without a teacher.
Schön's account of reflection, both reflection in (during) and on (after) action, and complemented in the revised learning cycle presented above by 'review' & 'explain'. Cowan's valuable third kind of reflection, 'reflection for action' (Cowan 1998), is a component of planning for the next action or of Kolb's active experimentation.

Schön rejects the idea that the work of professionals is adequately characterised by what he calls 'technical rationality', the (relatively unproblematic) application of specialist knowledge to specific problems and situations. Gibbons et al.'s account of the use of knowledge produced in Mode 1 is wholly consistent with 'technical rationality'. Technical rationality was certainly my dream as a new teacher and then teaching development enthusiast. A naïve early developer's view of improving teaching could be summed up as: "If we can only find out under what conditions people learn, we can apply our knowledge to our teaching, and all will be well." Schön rather sees the work of professionals as requiring continuing critical reflection to operate effectively under conditions of uncertainty and ambiguity, well beyond the guidance of rules. Nonaka suggests how critical reflection, reviewing and explaining, undertaken in social professional settings, together lead to advances in professional knowledge.

2.1.4 An account of learning revisited

So, bringing all the above together, we have learning, research and the production of knowledge as active and usually goal-directed activities, in which the intention to achieve some particular external goal and the intention to learn can comfortably co-exist. The learner, researcher or practitioner constructs, tests, refines, and extends their own knowledge and understanding, alone and in conversation. They progress through spirals of planning, action and reflection and explanation. This is the account of learning that will be used in this Overview.
2.2 Improving learning and teaching through accreditation

2.2.1 The initial question

Repeated consideration of questions 4, 5 and 6 from the Introduction, section 1.1, led me to a composite question - "Given the pressures on a lecturer's time - research, administration, teaching, service, work in their primary profession or in their professional body - what is the most cost effective way, institutionally and nationally, to increase the attention and priority given by an institution and a lecturer to their teaching and thus to increase the quality of teaching?"

These competing pressures have come, and still come, from a variety of sources. This is not the place for a detailed account of higher education policy initiatives over the last fifteen years, but an overview can be given. The Research Assessment Exercise (RAE) drives institutions to prioritise research for the large marginal return on effort it may provide. Student recruitment targets have to be met; financial penalties follow failure to do so. Universities may have targets for income generation through staff consultancy. Academic standards must be maintained, and at the same time retention targets met. External quality assurance requirements have imposed intermittent but urgent demands to improve teaching and, sometimes more important, to show improved teaching. Audit brings its rather more stately requirements to produce, implement and demonstrate the smooth functioning of well-developed and documented systems. The 2003 Higher Education White Paper alone brings with it The Higher Education Academy; training for all new higher education teachers from 2006; personal development portfolios for staff; participation increasing towards 50 per cent of those aged 18–30; further development of foundation degrees; top-up fees...

As well as competition with teaching for attention (RAE, income targets) there is also support for giving teaching greater priority. Some aspects of
quality assurance have done this. Strategies for learning and teaching and for human resources must be produced and implemented. The Teaching Quality Enhancement Fund and its National Teaching Fellowships, Fund for the Development of Teaching and Learning and earlier its Teaching and Learning Technology Programme; the work of the National Disability team and the projects that it supports; and the Centres for Excellence in Teaching and Learning; all of these offer substantial additional resources for the enhancement of teaching, and all require additional work and attention within Higher Education Institutions (HEIs) to apply for and then support. And there has been support for the development and accreditation of higher education teachers, through the Staff and Educational Development Association, the Institute for Learning and Teaching in Higher Education and now The Academy.

This (partial) list shows potential levers for improving teaching, and also competitors for attention. In this complex environment the need for developers to be ‘principled, pragmatic opportunists’ (Baume 2003:3 p 160 – Publication 6) remains evident.

Several possible alternative answers were considered to the question at the start of 2.1.1. Indeed I have adopted and followed each of these answers at some stage:

**Schemes or awards for excellence in teaching.** These are attractive because they embody a very simple model of learning – the view that, as Thorndike (1911: p 244) in his original formulation put it, "Of several responses made to the same situation, those which are accompanied or closely followed by satisfaction to the animal will, other things being equal, be more firmly connected with the situation, so that, when it recurs, they will be more likely to recur...". Rewarding excellence should lead to a repeat of excellence, and perhaps also to the imitation of that excellence by others. I developed one early scheme at London Guildhall University in 1994. Over time it became clear why such awards were not the most cost-effective way to improve teaching. Most academics owe their primary loyalty to their
discipline. Development work, to be effective, must work at least partly with and in the disciplines (Jenkins 1996). (Baume 2002:1 - Publication 4, pp 173-4, acknowledges the impact of Jenkins’ view on the need to work with the disciplines on my own development and practice.) A teaching award made to a lecturer in discipline x would be mostly likely to encourage other lecturers within discipline x, but might not have much influence on the practice of lecturers in disciplines y, z and the rest of the alphabet. Also, according to Gibbs et al. (2003, p 34) “The harsh reality of most recognition and reward schemes is that most applicants do not get rewarded – they get rejected.”

**Ensuring that teaching and research have parity of esteem in making academic promotions.** Dearing's work showed that, whatever promotions policies and criteria were in place, only 3% of teaching staff felt that teaching was valued in making promotions (NCIHE 1997). Attempts to influence university policy towards such parity of esteem met with little success – even in a teaching-led University, the idea of research retains a talismanic power.

**Providing good initial and continuing staff development and training.** This is indeed essential. But, to be effective, staff development must be embedded in a matrix of institutional, disciplinary and national culture, expectation and reward, and, as suggested below, qualification and accreditation. (Earlier in my career as a developer, I acted as if I believed that everyone would see the light if only I waved it about high and hard enough.)

**Writing and publishing about ways to improve teaching and learning in higher education.** Colleagues including Phil Race, Sally Brown and Graham Gibbs have clearly had a large effect through such work. I have contributed, and continue to contribute to, this literature (for example Baume and Baume 1992, 1996:1,2,3 and 1997:1,2; Baume 1996:1,2,3; 1998:2; 2000; 2001:2, 2003:3 – Publication 6; Baume 2004:2; Endean and Baume 2004). But 1991, when we began teacher accreditation, it seemed
that writing and publication on improving teaching and learning were well under way, and that adding lots more to the good and growing literature on learning and teaching would not give the best overall return on effort.

I have more recently focussed for the most part on writing about staff and educational development, seeking a multiplier effect, hopefully helping relatively small numbers of developers to work more effectively with much larger numbers of teaching staff with a correspondingly huge-scale impact on students' learning (Edwards and Baume 2003, Kahn and Baume 2003, Baume and Kahn 2004).

Several factors led me to Accreditation – to the idea that an appropriate national qualification in teaching in higher education, interpreted and implemented as a program of study within each HEI, was likely to be a very cost-effective way to improve teaching. Factors leading to this view included:

a. Experience in developing, and then supporting the implementation of, a national scheme to accredit playworkers. (Experience with playworkers also led me to discover the powerful role that underpinning values play in a professional qualification framework - Calibre Training 1990 (1-4).)

b. The view that students are led by their perceptions of demands made by assessment, and the extrapolation that the same might apply for lecturers.

c. The views of leaders of induction and training courses for new lecturers who explained that, in the absence of some institutional or national qualification or requirement, it was very difficult indeed to persuade either new staff or the institution as a whole to take participation in such courses seriously, other than a few enthusiasts.

My work on teacher accreditation is described in Baume and Baume (1996:4 – Publication 1) and Baume (2003:3 – Publication 6). Another account is given in UUK (2004). The scheme devised for the Biotechnology and Biological Sciences Research Council is described in BBSRC (2001).
Appendix 3 summarises work in the development and operation of accreditation frameworks and schemes. But it does not show the flows, the relationships, the learning. These flows and relationships are worth some attention, as they show coherence and development over some fifteen years of work.

2.2.2 Learning from and about accreditation – values

My work to develop an accreditation framework for playworkers generated some powerful learning. Wide consultation was clearly essential to ensure the quality and the acceptability of an accreditation scheme to those whom it is intended to accredit – the spoken views of some 200 playworkers informed the development of their accreditation scheme. Also, it became clear that the capabilities of playworkers could not wholly be represented by a series of NVQ-style competences. Something extra was needed. In a breakthrough moment during a consultation with playworkers it became apparent that this ‘something extra’ was orthogonal to the competences, underpinning or informing them. These came to be called the ‘underpinning principles and values’, more recently just ‘underpinning values’.

Such underpinning principles and values have been a feature of every accreditation scheme in whose development I have been involved. Underpinning values are vital to an adequate description of the capabilities of a professional in any field. In the early development and operation of SEDA’s Teacher Accreditation Scheme we received very strong opposition, from members of our advisory group, to the values: “You cannot specify another person’s values”. In response, we realised that it mattered less what they were called – values, principles, something else – than that they, in the phrase that came to be used, “demonstrably informed and underpinned” participants’ attainment of the outcomes.
2.2.3 Learning from and about accreditation

Two early schemes, for playworkers and the Institution of Environmental Health Officers, together provided vital experience in both the development and the detailed implementation of accreditation processes as well as frameworks. These experiences showed the importance of clarity and simplicity in accreditation - successive schemes have become progressively simpler. They also showed the need for examples, practice, conversations, support and feedback as well as formal scheme documents and handbooks.

As described in Baume and Baume (1996:4 – Publication 1), the SEDA Teacher Accreditation Scheme used seven underpinning values and eight outcomes. How was the learning reported in the Introduction about the importance of context to learning and to practice applied? By pushing the scheme away from prescribing a 'national curriculum' and a standard design of course. SEDA specified only the outcomes and the underpinning values, and left everything else to the individual institution. Ward and Dubos (1972) famously first suggested "Think globally, act locally". Simultaneous tight-loose coupling was described by Orton and Weick (1990) as a necessary quality for coherent yet flexible organisations. The key is to be tight and loose about the right things: National (and indeed international) standards, local implementation: Tight about outcomes and values and their assessment, loose about processes for attaining them. These precepts have proved an appropriate basis for the design of accreditation schemes in higher education, later adopted by ILTHE.

The SEDA Fellowship built closely on the approach adopted for teacher accreditation, with clear outcomes and underpinning values. It has been taken up by over 70 developers. SEDA Fellowship additionally requires an annual cycle, explicitly Kolbian, of continuing professional development. CPD involves writing a reflective report and gaining feedback on it from colleagues. A very similar approach was built into ALDAP and is informing current and future accreditation schemes.
2.2.4 Making teacher accreditation happen nationally

Early work on national teacher accreditation is described in Baume and Baume (1996:4 – Publication 1). More recent work is described in Baume (2003:3 - Publication 6) and summarised in Appendix 3.

2.3 Improving learning and teaching through portfolio assessment

Work on portfolio assessment is described in Baume and Yorke (2002 – Publication 9, 2004 – Publication 10). It directly underpins my work on accreditation. The course whose assessment by portfolio is reported is H851 – Teaching in Higher Education, a 30 M-level credit Open University course whose development I had chaired (led).

2.3.1 The importance of portfolio assessment

Assessment is the point at which accreditation decisions are made. Assessment means passing the course or not, gaining accredited teacher status or not, perhaps completing probation or not. Employer and lecturer alike care that assessment is properly conducted.

It has been proposed (Baume 1998:2) that good assessment is:

- Valid (assesses the attainment of the learning outcomes),
- Fair (a very difficult concept to work with as students and staff have strong and conflicting ideas about what makes assessment fair),
- Efficient (making the minimum necessary use of the time of the assessor), and
- Reliable (different assessor would give the same work the same mark or grade)
Co-researcher Mantz Yorke and I began with the list of virtues offered above. Validity of assessment on the course was good. It has been suggested elsewhere (Baume 2000:1) that the most valid possible assessment task is simply to say to the student "Show how you have attained the learning outcomes of the course." Assessment on the course we studied, which explicitly asks the course participant to make a reasoned and evidenced claim that he or she has attained the required outcomes underpinned by the required values, comes close to this ideal of validity.

Fairness was too difficult to find a research handle on.

The efficiency of assessment, determined in part by the time taken to assess each portfolio, had already been determined in the design of the course, and assessors paid accordingly. It was unlikely to be a fruitful research topic here.

Reliability is known to be often a major problem in assessment. For example Newstead and Dennis (1994) describe an experiment in which six psychology essays produced under examination conditions were each assessed by 14 experienced psychology assessors. The most reliably marked essay showed a difference between the lowest and the highest mark awarded of 13%. For the least reliably marked essay, the difference was 35%. There is no reason to suspect that Psychology is untypical in this regard, with obvious differences to be expected in assessment reliability between subjects lying along a numerical – verbal axis.

2.3.2 Reliability of and assessors' rationales for portfolio assessment and judgments

The first study (Baume and Yorke 2002 – Publication 9) found reliability of assessment to be towards the upper end of the published range of reliability data for portfolio assessment, certainly at the level of the individual outcome. (It would have been disappointing to have found otherwise, given
the huge efforts that The Open University applies to assessor briefing, training and co-ordination on all its courses.)

We explored in the paper the implications of our results for design of assessment schemes. I have later realised that these implications apply equally to the development of accreditation frameworks. The main implication is that an accreditation framework should have ‘as few elements as is consistent with requiring the teacher to show his or her capability in the key distinct areas of their work. These key areas may be planning, teaching, assessing, reviewing their work and continuing to develop as a teacher, an account later used in an accreditation framework and process for Open University tutors (Baume et al. 2005).

This analysis of assessment also led us to propose two further loops of learning and application. Tutors could direct teaching effort to the outcomes on which students did least well. And assessors could prioritise areas where assessor disagreement was greatest.

But assessment is not simply a statistical matter. Marks and numbers represent judgements. It was important to understand the judgement process itself. This wish to understand led to the experiment described in the second paper (Baume and Yorke 2004 - Publication 10). The experiment described in this paper in turn led to a more sophisticated process for improving assessment: The assessment criteria that were surfaced through the experimental process could themselves be subject to critical review by the course team and assessors together; consensus on criteria reached; and the revised criteria fed back into future rounds of teaching assessment, with care to avoid an upwards drift in standards.

2.3.3 An outer cycle of learning

Kolb presents a single cycle of learning from experience. I have extended Kolb’s account to describe ‘experiencing in order to learn’, that is in order to achieve particular learning outcomes, rather than ‘learning from experience’. 

David Baume
There is always the possibility of an outer loop, of a further cycle of learning or meta-learning. Argyris (1993) describes this outer loop, with anticipations of Nonaka (1994), as further learning conversations undertaken about some initial learning from experience. Baume and Yorke (2004 – Publication 10) offers an account of such outer loop learning applied to the operation of the course and its assessment. Flavell (1987) and others (Schoenfield 1978) describe this outer loop as meta-cognition; thinking about thinking; thinking about learning; planning, thinking and then reviewing the effectiveness of the thinking; applying a Kolbian learning cycle to the processes of thinking and learning.

The work on portfolio assessment is important for three reasons beyond improved accreditation practice.

First, it is important that lecturers joining the profession see assessment – their assessment – being done well, and have confidence in their assessment and hence in their accreditation.

Second, it is important that lecturers have the chance to explore what it means to be assessed well. (A further possible step, not yet taken: It would be desirable, after assessment, to take the new lecturers through their own process of being assessed. This would help them better to understand assessment. It would also help them to feed this learning forward into their own practice of assessing.)

Third, it is important that lecturers see a systematic process of the evidence-based improvement of assessment, again so that they can apply a systematic and evidence-based approach to their own improvement of assessment, teaching and other aspects of their work.
2.4 Improving learning and teaching through educational development projects

2.4.1 Projects compared with steady state

From the Kolbian, learning-oriented perspective adopted throughout this Overview, we may usefully contrast the project with what we may call 'steady state', such as the operation of a course year-on-year. Projects are typically characterised by defined start and end points; a defined project team; more or less clear goals; and plans of action. Of course, both 'project' and 'steady state' here are ideal types. Projects do not always have clear start and finish points, being planned and bid for, and hopefully influencing practice after the formal finish. And steady state is rarely all that steady; even an apparently steady state activity such as running a course should be subject to review, and thus to change.

There is a general shift in our working lives from steady state to a series of projects; to, in Vail’s phrase, evocative even for non-kayakers, "permanent white water" (Vail 1996). A benefit of this sometimes uncomfortable and disruptive shift from steady state to projects is that the project clearly and readily has a rhythm, a cycle, that can make Kolb's cycle of learning from experience more naturally explicit and hence productive.

2.4.2 Projects and their outcomes

Much of my recent research and development work on projects has involved the adjacent twins of specifying project outcomes and evaluating projects' attainment of their outcomes. Work on project outcomes (as for example in Baume 2004:1) is in direct line of succession to my very earliest academic publication, Baume and Jones (1973). This was concerned with the specification of educational outcomes (then more commonly called objectives) for courses. The academic root of these two pieces of work,
done 30 years apart, is substantially the same. The first question about projects - "What are you trying to achieve? What are the intended outcomes, effects, of your project?" is almost identical to the question, first to myself as a lecturer and then over time to colleagues and to the wider higher education community - "What are you trying to achieve? What are the intended outcomes, effects, of your programme, your teaching?" I seem to have known for a very long time that "If you don't know where you're going, any bus will do" (Anon). Without explicit clarity of purpose, of intended outcome, it is difficult to:

- Plan (whether a project or a course) with rationality and confidence
- Know - whether 'know' here means 'assess' or 'evaluate' - that the project or course has been successful

Projects and courses have been seen to founder in this pit of unclarity, often without anyone involved understanding what was wrong.

2.4.3 Outcomes and evaluation

Note the synergistic relation between on the one hand intended outcomes, whether for a project or for teaching, and on the other evaluation (of a project) or assessment (of learning). Without evaluation or assessment, we cannot know if our project or our teaching is succeeding. This relationship works in the other direction as well. Without evaluable intended outcomes, evaluation is very problematic. With both evaluable outcomes and the act of evaluation, it is possible to learn about what is working and what isn't, and why, and how. It becomes possible to understand and thus improve the effectiveness of what we do, whether we are running a project or teaching. In Kolbian terms, 'review' can mean both, in project language, 'monitor' (during the project) and 'evaluate' (at the end of the project).

Much of my work uses the refined version of Kolb's learning cycle shown earlier – plan, do, review, seek to explain, use the emergent explanation to
inform future planning and action. The account here of the importance of outcomes requires a slight elaboration of the refined model.

Planning now has two stages rather than the one implied in the model. First, planning the intended outcomes, and then, planning what actions will be required to achieve them.

After completing the project or the teaching, there is still only one review question, but it is more complex: "Why have the project's activities or the teaching led, to the extent that they have, to the intended outcomes being achieved?"

This all offers a very pure approach. First the intended outcomes, then determination of appropriate activities to achieve these, then reviewing actual outcomes against goals. Life is not always so pure. Sometimes we start with a clear idea of what we want to do rather than of what we want to achieve. What matters is not the sequence of planning steps, but the eventual consistency – in Biggs' (1999) highly appropriate phrase, the 'constructive alignment' – between intended outcomes and intended activities to achieve these outcomes.

2.4.4 Learning from a group of projects

Work on learning from a group of projects is described in Baume (2002:1) – Publication 4.

2.4.5 A further example of learning from a group of projects - IT Term at Oxford Brookes University

Work on IT Term at Oxford Brookes University is described in Baume et al. (1996 – Publication 2). This paper is a summary of a much longer systematic evaluation report (Baume 1996:3) of the learning from and about each project within IT Term and the term as a whole.
2.4.6 Learning from experience, intending to learn

A super-ordinate aim, in any professional or purpose-oriented setting where an activity is undertaken more than once, is to improve the methods used to achieve the goals, the intended outcomes, of the activity. Even, perhaps, to improve the intended outcomes themselves. ('Improving the intended outcomes' might mean making the outcomes better attuned to the larger goals of the organisation, or to the goals of the various stakeholders in the activity - for example clients, funders or relevant professional bodies.) Straight Kolbian learning from experience may thus occupy a much smaller role, have much less importance, certainly in education, than was previously thought. Purposive learning from experience, whether that purpose is to achieve specific learning outcomes or to improve the educational systems and processes, should probably be given much more prominence.

2.5 Improving learning and teaching through evaluation

Evaluation is closely associated with the Reflective Observation and Abstract Conceptualisation stages of the Kolb cycle, that is with the Review and Explain stages of the revised cycle. There are also similarities between evaluation and assessment; in both cases we ask what we know, and then ask how we know it.

Evaluation offers an outer loop around all of the work described in this Overview. Evaluation would enable us systematically to investigate whether accreditation is in fact a cost-effective way at a national level to improve teaching and student learning (Section 2.2). The research into the assessment of portfolios (Section 2.3) evaluates the assessment process used on the course. The work on projects (Section 2.4) stresses evaluating the extent to which projects have achieved their stated outcomes as well as extracting fresh understanding on the management of the current project and projects in general. Evaluation seeks to improve a project, to make it accountable to stakeholders, and to understand the project, its processes.
and their effects. It would be valuable to evaluate evaluations themselves, to see how far they had generated understanding of current projects, and led to the improvement of future projects – and, of course, to improvement of the practice of evaluation.

2.5.1 How can we know? – Disciplinary approaches to knowledge

I have recently come to realise that behind all of my questions lay the same meta-question - "How can we know?" Some personal academic background is required here. My first discipline was Physics. In studying Physics I had absorbed, as well as theories and models and facts and ways of reasoning, a very high standard of proof as a *sine qua non* of valid knowledge and theory. Later, studying Philosophy of Science, I learned to articulate the nature of the very high standard of proof I had previously encountered in Physics. A theory, I learned, can never safely be considered to be proven whilst any condition where the theory might be contradicted remains unexplored. Any theory worth its salt entails an infinite number of particular predictions. One, just one, instance of a prediction of the theory being confounded by an observation means that the theory is disproved. Thus science advances by inventing, testing and (usually) refuting theories, and then producing the next theory, and so on. A cyclical process of the production of scientific knowledge, with similarities to Kolb's account of learning: This, I was taught, was how science worked and advanced. (See for example Popper, 1969)

I initially tried to bring the physical sciences model and rigour to the investigation of student learning. It rapidly became clear that this was an impossible quest. Words including 'hypothesis', 'prediction', even 'truth', had different meanings in education than they did in Physics. (This was happening in the 1970s, before post-modern thinking had acquired much currency.) Research into education had different goals, methods and standards from research into Physics. It sought useful models or
representations. The theories it talked about would not have been recognisable to a physicist as theories. Research into education sought understanding rather than proof, correlation and probability rather than certainty. Good and productive questions were valued almost as highly as good answers. I had to find new ways of knowing about the relations between, for example, teachers' actions and students' learning. I became and remain intrigued by the relationships between research and evaluation. This whole Overview is about research and its application and impact. The focus here is on evaluation.

2.5.2 Evaluating staff and educational development

In 1993 with Carole Baume I researched UK staff and educational developers' practice in evaluation (Baume and Baume 1995 – Publication 7). A conclusion was that, "while evaluation of staff and educational development events and programmes was widespread, much less evaluation was carried out in respect of policy." Ten years on, most of the evaluation of staff development alas still involves little more than gaining immediate feedback on participant satisfaction using the (in)famous happy sheets. Any half-way competent developer can get very positive responses to these without undue difficulty.

Baume and Baume (ibid.) describes an evaluation questionnaire that encourages participants before a workshop to identify what, in addition to the workshop's intended outcomes, participants want to get out of the workshop. After the workshop, the same questionnaire asks them to say how far they got from the workshop what they wanted as well as what the workshop leader had wanted them to get. Baume (2001:4) describes one benefit of this approach – the workshop can be very responsive to participants' expressed needs – and also one drawback – the risk of setting up expectations that cannot be met in the time or with the resources available. Baume (2003:1 – Publication 5) introduces some useful new
elements to the evaluation of staff and educational development as well as reviewing older methods.

2.5.3 Outcomes and evaluation

More recently, something implicit in the workshop evaluation tool described in 2.4.2 has become much more explicit – the realisation that there is one overarching evaluation issue and question: "How far have the goals of the project been met?" The idea of 'evaluation against intended outcomes' is used in a number of ways. I have helped ESECT (2004), a £0.7m 30 month HEFCE-funded project on student employability to which I am external evaluator, to clarify its intended outcomes, from its initial statement of intent to 'make a decisive difference' into a rather more detailed but still concise and clear form which is currently guiding the operation of the project and the production of its many deliverables. I have done the same for the two FDTL Phase 4 projects for which I am also external evaluator.

Recent work on outcomes and evaluation has built on the distinction made in Baume (2003:1 – Publication 5) between project outputs or deliverables, things that will be produced, and outcomes, particular differences that will be made. This work has also extended the idea of searching for plausible and evaluable proxies for any outcomes which are not directly attainable or measurable within the lifetime of the project. Hall and Louks' (1978) seven-step account of a project's impact (an eighth step of course is ignorance) has been modified. An essential further stage (planning to act) has been added. The account has then been stripped back to three essential and evaluable stages of project impact – awareness, planning to act and action.

It has been objected that, in complex settings, it is impossible to disentangle the web of cause and effect. The idea of cascade goals attempts to address this. Projects are encouraged to identify their final intended outcomes. In the case of ESECT, the outcome is a set of student behaviours, namely that students develop, define, work towards, attain and review their attainment
of their employability goals. It is then easy to work out goals for those who work directly with students – that they support students to undertake a piece of student behaviour, namely that students develop ... and etc. as above. And then again goals for those who work indirectly with students – that they create the environment and policy and resourcing framework within which students are supported to develop... and etc.

2.5.4 Supporting project evaluation

In 1997 Carole Baume, Graham Gibbs and I won the contract from the HEFCE to coordinate the Fund for the Development of Teaching and Learning (FDTL), the first phase of which was then under way. The contract was later extended to embrace, first the Teaching and Learning Technology Programme (TLTP) and then the support of the National Teaching Fellowships scheme, all under the umbrella of HEFCE’s Teaching Quality Enhancement Fund (TQEF). I defined, planned and edited a series of Project Briefings, covering (currently) 20 topics (TQEF NCT 1997). The Project Briefings were intended to capture best practice in concise (2-4000 words) and accessible form, with examples – they are firmly ‘how to’ guides.

The first draft of the Project Briefing on Monitoring and Evaluation (ibid) made explicit and referenced use of Baume and Baume (1995 – Publication 7). This draft briefing was used as the basis of a workshop. Participant feedback on the draft, and suggestions for improvements and additions, were fed into the next edition. The revised edition now forms part of the TQEF Project Manager’s Handbook. The process of writing the Briefing itself thus used a simplified version of Kolb’s learning cycle. (There was little need to conceptualise or explain why the changes were being requested - the reasons were mostly obvious.)
2.6 Kolb and intentional learning

The work and this Overview extend Kolb’s model of learning as a cyclical process to describe and plan intentional, rather than serendipitous, learning, in four overlapping contexts – accreditation, portfolio assessment, projects and evaluation. This section summarises the relationships between the work in each area and the model; summarises the impact of the work; and then considers Kolb and intentional learning in rather more detail.

2.6.1 Accreditation and intentional learning

Development and accreditation involve a formal process of directed learning, from experience and / or from being taught, followed by assessment against a defined professional standard. Accreditation takes two distinct forms for two different groups of users:

Experienced teachers review and explain both their teaching and their learning about their teaching. They review and explain explicitly in terms of the capabilities, underpinning values and underpinning knowledge specified by the accreditation framework. Some of the process is conducted in conversation. Accreditation is thus more than a simple recognition of past learning – it is also a learning process in itself. The accreditation is developmental, more like Nonaka’s than Kolb’s account.

New teachers use the framework of specified capabilities, underpinning values and underpinning knowledge to inform both their work as a new teacher and their learning from their current experience. We thus have outcome-directed experiential learning. The teaching and the structured reflection thereon by new teachers are, as for experienced teachers, developmental.
2.6.2 Portfolio Assessment and intentional, outcomes-specified learning

The assessment of portfolios is a vital part of the accreditation process. This has been researched very closely in the work described here. Portfolio assessment is the part of accreditation where, in Kolbian terms, the teacher shows in their portfolio the whole of the learning process through which they have gone, all four stages of Kolb’s cycle, as well as showing the outcomes they have achieved through this work. The teacher informs each of these stages by theory and by accounts of the practice of others. And the teacher shows how he or she has achieved the intended learning outcomes of the accreditation framework, together with its underpinning knowledge and professional values.

Assessment is the crunch point, the point at which (in the courses studied) the teacher makes the explicit, evidence-backed claim that they have achieved the course outcomes and at which the assessor judges whether or not they agree. The research reported in Baume and Yorke (2004) has identified, in great detail, how these judgements are made, and has identified a process – again based on Kolb – for improving the reliability of these judgements and the teaching of the course.

2.6.3 Educational Development Projects and purposive, outcomes-specified action

Educational development projects are also planned, run, monitored and evaluated as goal-directed undertakings involving cycles of learning. The chapter on learning from educational development projects (Baume 2002 – Publication 4) shows the use of one arc of the Kolb cycle in action, a sector that we may call purposeful review, explanation and - not planning, but rather providing tools and ideas to be used by people planning educational development projects. The chapter thus adapts the work of others for a new purpose. Kolb’s cycle of learning and the account of purposeful learning from
experience are thus taken a step further in this work on educational development projects.

2.6.4 Evaluation of purposive, outcomes-specified activity

Evaluation, like assessment, involves making judgements against criteria. In direct parallel with the outcomes-oriented experiential learning which underpins the work in accreditation and in assessment, described earlier, I have researched, developed and been invited to apply to educational development projects a form of outcomes-oriented evaluation. This evaluation asks, at the start of a project, “What particular differences do you intend your project to make?” During / at the end of the project it requests, “How is your project leading / has your project led to these outcomes being met?” Again the emphasis is on goal-oriented behaviour. Again a goal-oriented learning cycle is used, this time the project being conducted in part as a learning cycle, with goals set, appropriate activities planned and undertaken and deliverables delivered, and attainment of outcomes reviewed.

Projects are invited to ask two further questions – “Why is your project leading to these outcomes being met?” and beyond that “What makes you think so?” Asking these two questions involves digging for evidence- and reason-based interpretation and theory emerging from the project. Thus monitoring and evaluation become scholarly activities as well as activities to deliver accountability and improvement. The project then re-plans its activities, its deliverables, and maybe even its intended outcomes, to use what it has learned. The learning from the Kolbian cycle around the project can also be made available to other projects, as throughout Baume 2002 (Publication 4).
2.7 Originality and impact

2.7.1 Impact of work on accreditation

From the launch of the scheme in April 1993 at the then Polytechnic of North London (which ceremony included the presentation of the first accredited teacher course certificate to Danny Saunders of the then Polytechnic of Wales by Sir Geoffrey Holland, then Permanent Secretary at the Department for Education) until the SEDA Teacher Accreditation Scheme stopped accepting new applicants in 2002 (in deference to ILTHE), 65 programmes had been recognised and 3100 teachers accredited by SEDA. The further influence of the work described and analysed here on teacher accreditation has been considerable.

The form of the SEDA Teacher Accreditation scheme – simple and clear outcomes with explicit underpinning principles and values – is manifest in the ILTHE accreditation framework, to which I made a substantial contribution. In the five years of the ILTHE's existence, before it is absorbed into the new Higher Education Academy, ILTHE has recognised some 16,700 lecturers. 76% of these have been accepted into membership by direct entry, and the remainder by graduating from 147 ILTHE-accredited programmes in 114 HEIs. Thirty-nine of these programmes were recognised en masse by ILTHE on the basis of their prior recognition by SEDA, at a recognition event in which a colleague and I represented SEDA. This recognition again demonstrates the close relationship between the ILTHE and the SEDA standards.

The further impact of my work is shown in my being invited to develop other accreditation frameworks as summarised in Appendix 3.

I have recently been invited to design and co-lead consultation events for the Higher Education Academy and to undertake preliminary research into occupational standards. Both these pieces of work are intended to inform the development of the new standards which the 2003 White Paper (DFES
2003) requires, evidence of the continuing and future impact of my work on the accreditation of teachers in higher education.

The story of the development of higher education teacher accreditation is also told in UUK (2004).

2.7.2 Impact of work on Portfolio Assessment

On the basis of my work on portfolio assessment I was invited to:

- Write briefing papers on the assessment of portfolios and on helping people to prepare portfolios for the LTSN Generic Centre (Baume 2001:2, 2003:2)
- Give a paper on the topic to an LTSN conference.
- Give a paper to the AAHE assessment conference in Charlotte NC
- Write a paper for ERIC (Baume and Yorke 2000)
- Write a case study for a book on assessment (Baume and Yorke 2001 – Publication 8)
- Write an article for the Times Higher on developing a teaching portfolio (Baume 2001:1)

Citations of this and other work are shown in Appendix 1.

2.7.3 Impact of work on Educational Development Projects

On the basis of the chapter on educational development projects seen in early draft, I was invited to run a session for the 2002 Conference of the International Consortium for Education Development in Perth, WA, in which I outlined, and then invited participants to test and apply in their own settings, my main conclusions about the management of educational development projects. This invitation speaks to the impact of the chapter even ahead of its publication. The workshop design, featuring me reviewing
and explaining and then participants planning how to do, shows Kolb informing my work yet again.

2.7.4 Impact of work on Evaluation

The Project Briefing on Monitoring and Evaluation has been used by all four phases to date of FDTL, 130 projects with a total budget of £27m. It was also used by 65 TLTP phase 2 and 3 projects with a total budget of £22.75m. It will further be used, in a modified form, by the FDTL 5 phase projects starting in 2004 with a further budget of £7m. It is also used by 80 National Teaching Fellowship Scheme (NTFS 2004) winners from 2000-2003 with a total budget for their development projects of £4m and will be used by the 50 2004 NTFS winners with a further budget of £2.5m. For FDTL 5 I have produced a new Project Briefing on Goal Setting and Clarification and run workshops on Goal Setting and on Monitoring and Evaluation.

A modified version of the Project Briefing on Monitoring and Evaluation was used by the National Disability Team (NDT) (NatDisTeam 2004) which now coordinates 54 higher education projects with a total budget of £5.5m. In late 2003 I ran a workshop for NDT’s second phase projects on goal clarification and project monitoring and evaluation, produced a Project Briefing on Goal Clarification and further revised their briefing on monitoring and evaluation.

In sum, my Project Briefings on Project Monitoring and Evaluation, with the workshops based on them, have directly influenced or are currently influencing some 330 educational development projects with a value of over £60m, and the work of the thousand or more project staff.

Current research, development, practice and reflection on evaluation will inform a further paper on educational development goals and evaluation to be submitted for publication in 2005.
2.8 Purposive learning from experience

Kolb describes a process of learning from experience where learning is almost a by-product of action and review and explanation, albeit a welcome by-product. Kolb’s account has here been extended in this Overview to describe intentional learning from experience. Learning is reconceptualised in terms of the explicit intention which the learner brings to his or her movement round the cycle: Experiencing to learn rather than learning from experience. Teaching, the supporting of learning, is further reconceptualised as designing cycles of experience, review and explanation, and then supporting learners round these cycles. Section 2.4.2 showed two linked kinds of planning – of intended learning outcomes, and of the activities or methods or projects that will be used to achieve them. In the relationships between the intended learning outcomes and the activities we use to help students to attain these outcomes lie all of our theories and models, explicit and more often tacit, about learning and about teaching. The discussion of accreditation, and earlier accounts of Nonaka’s analysis of organisational learning, describe a process of surfacing tacit knowledge and making it available for review and testing and further development, an approach used extensively in this work.

‘Learning outcomes’ need further deconstruction. Learning outcomes can exist at a number of levels – minimally, at programme, course, class and perhaps individual learning activity level. They will have been articulated with various degrees of clarity at each of these levels. In general, the more explicit these intended learning outcomes are, at each level, the better they can serve to guide teachers and learners in planning, undertaking, reviewing and explaining learning – in going round the learning cycle.

Schön (1982) talks of reflection on action, i.e. reflection after the event, but much more about reflection in action. Cowan (1998) very valuably adds to this list reflection for action, reflecting on and learning from prior experience in order to inform the planning of future activity. Reflection in action is, Schön suggests, an essential characteristic of the work of a professional.
The need to reflect in action marks the important and difficult point where the professional ventures beyond the rules-governed behaviour of the technician and strides into the swamps of complexity and uncertainty that are the professional's working environment.

The difficulty is that reflection in action is very difficult indeed to locate, to identify, to see, as Eraut (1995) has commented.

How to resolve this difficulty?

Learning cycles have some of the qualities of fractals (Mandelbrot 1982). That is, they exist, they occur, they are employed, in essentially the same form that Kolb describe, at a wide range of scales from micro – lasting a few seconds – to macro – lasting a lifetime, and at many scales in between.

Examples: A child learns about the inadvisability of touching lit light bulbs in under a second, a piece of learning that mostly stays with the child, growing in depth and understanding as the years pass. By contrast, some longer learning cycles, mainly about ideas and also about people, may remain unclosed in a lifetime.

So, instead of looking for something elusive – reflection in action – why not look for tiny cycles, intersecting micro-cycles if you like, of learning, of reflections on and for action? Developers and teachers may see these in their own work, where they adjust the schedule of a workshop or class on the fly in response to their perceptions of participants' degree of engagement and particular direction of interest. Musicians report doing something very similar during performance (Springsteen 2003).

This interest in micro-scale reflection and learning is evident in the experimental work with Mantz Yorke on the assessment of portfolios (Baume and Yorke 2004 – Publication 10). Assessors were asked to identify, line by line, as they assessed, why they were making the assessment judgements that they were making. As well as leading to local learning by each assessor, that work also helped the development of a further, year-long cycle in which the course itself may legitimately be said to learn from the analysis of
assessment results, and then to prioritise work to overcome students’ difficulties with the course and work by staff to improve assessment. The idea of a course itself learning is also described, explicitly, in Hipwell and Baume (1976).

Kolb and Nonaka together provide a tool that a lecturer (or any professional) can use to interrogate and improve their own practice. The lecturer talks about their teaching; makes and tests new sense of it; finds what they need to change in their teaching; and then plans, undertakes, reviews and explains the effects of these changes. Developers can assist this change process, in several ways, as well as applying it to their own work as developers. We can steer the process through the use of a national accreditation framework and perhaps in the future through institutional and disciplinary interpretations thereof. We can offer additional questions and tools, for reflection and planning and doing and reviewing and explaining and re-planning. We can offer models and theories of teaching and learning, models which the lecturer can review in terms of their utility.

Thus the version of Kolb offered here which is based on action with the intention to learn, taken together with Nonaka’s emphasis on the social nature of knowledge production, offers a shared tool for joint use by lecturers and developers. It offers an heuristic that prompts and supports the development, and in due course the accreditation and continuing development, of each lecturer (and developer). It offers a tool for active learning by each lecturer, albeit in a social setting. The approach described here respects each lecturer’s motivation and current capabilities. It acknowledges the institutional and disciplinary contexts in which they work. At the same time, the approach is directed towards a common, though locally tailored, account and standard of teaching competence. This approach surfaces, and thus lays open to critical review and further development, the often secret scholarship of teaching.

Such an individualised and reflective approach towards a common standard of teaching is likely to deliver the requirement that, in the words of my oral
evidence to the Dearing Committee and of the 2003 White Paper on Higher Education:

“All students are entitled to be taught well...”

This single short statement describes the motivation for all of my work described in this Overview, indeed for most of my work in staff and educational development over the last fifteen years, and hopefully for at least some of my teaching before that.
References

Anon (undated), variously attributed as Spanish or Latin-American


Publication 1


Publication 2


Publication 3


Publication 4


Publication 5


Publication 6


Publication 7


David Baume


Baume, D. and Yorke, M. with Coffey, M. (2004). What is happening when we assess, and how can we use our understanding of this to improve assessment? *Assessment and Evaluation in Higher Education* 29 (4): 451-477. ISSN 0260-2938

BBSRC (2001) *Training and Accreditation Programme for Postgraduate Supervisors (TAPPS): An explanatory*
booklet URL: http://www.iah.bbsrc.ac.uk/TAPPS/TAPPS-explanatory%20booklet.htm, accessed 20th January 2004


NTFS (2004). *National Teaching Fellowships Scheme,* URL: www.ntfs.ac.uk accessed 24th January 2005


TQEF National Coordination Team (1997). *Project Manager's Handbook.* Milton Keynes, National Coordination Team - The


The papers numbered 1 – 10 are included in this submission.
4 The Evidence


Publication 10 Baume, D. and Yorke, M. with Coffey M. (2004). What is happening when we assess, and how can we use our understanding of this to improve assessment? *Assessment and Evaluation in Higher Education* 29(4): 451-477. ISSN 0260-2938
Tabulation of overview sections and themes against publications

<table>
<thead>
<tr>
<th>Section</th>
<th>Theme</th>
<th>Publications</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.2</td>
<td>Accreditation</td>
<td>1, 3, 6</td>
</tr>
<tr>
<td>2.3</td>
<td>Portfolio assessment</td>
<td>3, 8, 9, 10</td>
</tr>
<tr>
<td>2.4</td>
<td>Educational development projects</td>
<td>2, 4</td>
</tr>
<tr>
<td>2.5</td>
<td>Evaluation</td>
<td>2, 5, 7</td>
</tr>
</tbody>
</table>

A review of the development and first few years of operation of the SEDA Teacher Accreditation Scheme in its national and educational context, with analyses of current and future developments of the scheme and of critical success factors.
A national scheme to develop and accredit university teachers

Carole Baume, Oxford Brookes University, England, and Co-Chair of SEDA; and
David Baume, The Open University, England (from February 1997), and former Chair of SEDA

ABSTRACT

In 1991 the UK Staff and Educational Development Association (SEDA) developed and launched a national scheme for the training and qualification of teachers in higher education. Teachers are accredited if they demonstrate eight specified objectives underpinned by seven specified values. Programmes of staff development which assess teachers’ attainment of these objectives and values are recognized by SEDA. By the start of November 1996, SEDA had recognized 30 programmes and a similar number of programmes were at various stages of preparation for recognition. More than seven hundred teachers are undertaking recognized programmes or are already accredited. The paper describes the development and operation of the scheme; starts to locate the scheme in a wider framework of competence, preparation and accreditation of all staff who teach in higher education; and explores some of the operational, educational and philosophical issues faced by the scheme.

The context for the initial development of the Teacher Accreditation Scheme

Two major forces—a new emphasis on quality and accountability, and the emerging movement towards national vocational qualifications—strongly affected the environment in which the Teacher Accreditation Scheme was developed.

In 1990, UK Higher Education was still a binary system, with both universities and polytechnics. Universities were being called to account for the quality of their courses much more strongly than had previously been the case. A new system of Quality Audit for the universities was developed and introduced in October 1990 by the Committee of Vice-Chancellors and Principals (CVCP). Its first term of reference was to ‘consider and review the Universities’ mechanisms for monitoring and promoting the academic standards which are necessary for achieving their stated aims and objectives.’ (CVCP, 1992)

When the polytechnics became universities in 1992, they too were subject to audit. Many of them felt more comfortable with audit than did the older universities. The polytechnics had previously been subject to external scrutiny—of their courses, during the 1970s and 1980s, and later of their quality assurance arrangements—by the Council for National Academic Awards, which awarded degrees gained in the polytechnics. Universities had had no such external scrutiny. A thoughtful account of quality issues in UK higher education in relation to Quality Audit is provided by Barnett (1992).

Another quality system followed quickly on Quality Audit: ‘Section 70 of the Further and Higher Education Act of 1992 places on the HEFCE (Higher Learning Funding Council for England) the statutory obligation that it shall “secure that provision is made for assessing the quality of education provided in institutions for whose activities it provides, or is considering providing, financial support…”’ (HEFCE 1993). The HEFCE method is called Quality Assessment, and involves self and peer assessment of the provision of education. It is undertaken on a subject basis and includes classroom observation of teaching.

51
In addition to these two quality systems, professional bodies in many disciplines impose their own requirements on the design and operation of courses in universities.

In the early 1990s, the second major force to influence the higher education environment, the National Vocational Qualifications movement, was gathering pace. Although education had not yet been brought into this national vocational framework, it was clearly possible that education would in due course follow other professions. A further consideration was the growth of European professional qualifications.

The development of the Scheme

The development of the Teacher Accreditation Scheme began in November 1990. Development was initiated within SCED, the Standing Conference on Educational Development, one of the two predecessor organizations of SEDA, the UK Staff and Educational Development Association (SEDA, current). Audit had just been launched; Assessment was yet to come. Increasing pressures were being imposed, year on year, on the funding of higher education. And, as suggested above, the quality of university teaching was becoming an issue of concern.

A SCED working group had been established to see if there was scope for polytechnics and universities in the London area to share at least some of their induction programmes for new teaching staff. Logistical and organizational problems, and differences of institutional culture and style, made such sharing impractical.

However, the group went on to ask what turned out to be a much more fruitful question - What should be the outcomes of any induction programme for new lecturers? Two sets of answers emerged: new lecturers should know about their institution and about their roles in it; and new lecturers should have attained some definable level of competence in teaching. On the first issue the working party felt that goals and methods of induction are best determined within each institution. On the second issue, the working party felt some consensus might be possible.

Given that some consensus might be achieved, within the working group and perhaps beyond, on a standard of competence in teaching for new university teachers, what role might SCED play?

Influenced by the UK national movement to develop vocational qualifications, the working group considered offering individual accreditation to teachers. This felt an unwieldy process, beyond the reach of a small voluntary professional association. An alternative approach was that SCED could recognize courses of training for new university teachers which required participants to show that they had attained the specified competences. Teachers successfully completing these courses could then be accredited. The working party accordingly redefined its terms of reference as:

1 To consider the needs of the beginning teacher in higher education, and to express these as an initial training programme; and
2 To design a method for the accreditation of programmes between institutions.

(Baume, 1992, p. 8)

Development of the Scheme continued during 1991 and 1992. An important early step was the decision, or perhaps the realization, that competences or objectives alone were not sufficient to describe the work of a teacher, and that values and principles must underpin competences.

In Spring 1992, a pilot began with eight institutions, both pre- and post-1992 universities. Development culminated in a well-attended launch in April 1993 at which the first certificates of recognition for programmes and the first certificates of accreditation for lecturers were presented.

Outline of the Teacher Accreditation Scheme

The description that follows is adapted from the current Scheme leaflet. (SEDA, 1996a)

The Scheme recognizes programmes of training and development for teachers in higher education. Teachers who successfully complete recognized programmes are accredited by the Scheme.

A programme is recognized if it:

- requires teachers to demonstrate the achievement of each of the eight specified objectives of the Scheme, in a way which reflects each of the seven specified values of the Scheme;
- involves an appropriate mix of self-, peer- and tutor-assessment;
A NATIONAL SCHEME TO DEVELOP AND ACCREDIT UNIVERSITY TEACHERS

is externally examined and/or moderated; has a procedure for dealing with appeals against accreditation decisions; has a regular reviewing mechanism.

The eight objectives that each accredited teacher must demonstrate are designed so that they can be adapted to particular teaching roles and particular institutional missions. For accreditation, teachers should have:

1. Designed a teaching programme from a course outline, document or syllabus;
2. Used a wide and appropriate range of teaching and learning methods effectively and efficiently, to work with large groups, small groups and one-to-one.
3. Provided support to students on academic and pastoral matters;
4. Used a wide range of assessment techniques to assess student work and to enable students to monitor their own progress;
5. Used a range of self-, peer- and student-monitoring and evaluation techniques;
6. Performed effectively the teaching support and academic administrative tasks involved in their teaching, in their department, in their institution;
7. Developed personal and professional coping strategies within the constraints and opportunities of their institutional setting;
8. Reflected on their own personal and professional practice and development, assessed their future needs, and made a plan for their continuing professional development.

For accreditation, teachers must also show how each of the following values and principles underpins their work:

1. How students learn
   All teaching, academic administration and pedagogic research should be informed by an active searching out of a better understanding of how students learn.
2. Individual difference
   Helping students to learn must begin with a recognition that each student has individual learning needs, and brings individual knowledge and resources to the learning process.
3. Development
   Education is about the development of students' existing skills, knowledge and attitudes, and their confidence in themselves, so that they can take responsibility for their own learning. Our work with students should therefore empower and enable them to develop greater capability and competence for their personal and professional lives.
4. Scholarship
   At the base of a teacher's competence is an awareness and acknowledgement of the ideas and theories of others. All teaching should be underpinned by a searching out of new knowledge and a passing on to students of a questioning and analytical approach.
5. Collaborative working
   Much of our work as teachers is carried out as part of a team made up of teaching staff and academic support staff. The colleagueship and support of peers is as important as individual academic excellence.
6. Equal opportunities
   Everything a teacher does should be informed by Equal Opportunities legislation, policy and best practice. This requires an understanding of equal opportunities in the curriculum and in the institution.
7. Reflection
   Teachers, like all other professionals, do more than have and use competences. They also reflect on their intentions and their actions, and on the effects of their actions. They try to understand the reasons for what they see, and for the effects of their actions. They thus continue to develop their understanding and practice, and therefore improve their own learning.

Take-up of the Scheme

By the start of November 1996, 30 programmes had been recognized. Eight of these were in pre-1992 UK universities (including one specialist programme for teachers of architecture), four were in UK colleges of higher education (including one for teachers of speech and drama), one was in a polytechnic in Singapore and the remaining 17 were in post-1992 UK universities.

Of the further 27 programmes preparing for recognition at the start of November 1996, six were in pre-1992 UK universities, six in colleges of higher education and similar institutions, five in specialist programmes to train higher education
Involvement with the Scheme is thus spreading across the UK higher education sector. Interest and involvement from outside UK are growing. An up-to-date account of registrations and recognitions is available on SEDA's World Wide Web site (SEDA current).

Quality assurance for the Scheme

The Accreditation Steering Committee was set up in Autumn 1992 to establish, guide, support, advise and review the standards and operation of all SEDA Accreditation Schemes. It meets twice each year. It is made up of representatives of higher education and related organizations such as quality and funding bodies, trade unions and professional organizations; individuals present in a personal capacity, such as eminent professors of higher education; and SEDA officers. It is chaired by a university vice-chancellor.

In November 1995, the Steering Committee commissioned a review of the procedures and working of the Scheme. This review was conducted by Associate Professor Jackie Lublin of the University of Sydney, who surveyed a sample of six course leaders of recognized programmes, members of SEDA's administration and members of the SEDA Teacher Accreditation Committee. In her report she concluded:

‘In general the Scheme appears to be working well, and there seems no reason why SEDA would at this stage wish to rethink its approach to implementation.’

(Lublin, 1996(a), p. 9 and 1996(b))

Her report did point to several areas in which a review of procedures and guidelines would be useful. The Lublin Report is being considered by the SEDA Teacher Accreditation Committee at a retreat; some of the recommendations have already been implemented. In 1988, the Scheme will be subject to a larger external review which will assess the nature and effectiveness of the Scheme, as well as its operation.

The great majority of SEDA-recognized courses are also validated by the university in which they are run, and are thus subject to the quality assurance procedures of the university. In every case, the programme must also be subject to scrutiny by an external examiner or moderator appointed by the university. The Teacher Accreditation Committee is currently exploring ways in which these external examiners or moderators can be helped to ensure programme quality with respect to SEDA objectives and values, as well as with respect to the requirements of the university.

Links to other SEDA accreditation schemes

Once the Teacher Accreditation Scheme had been established, there was a demand for a similar process of development and qualification for the leaders of programmes for new university teachers, and then more widely for staff and educational developers in general. The Fellowships Scheme was launched in May 1994 (SEDA, 1995a; SEDA, 1996b). It also is based on defined objectives and values, this time with the addition of specialist topics, chosen by the developer, in which the developer is required to show particular staff and educational development expertise. By November 1996, 18 Fellowships had been awarded and 23 candidates were preparing for accreditation. An Associate Fellowship, for new, part-time or specialist staff and educational developers, has recently been launched, with a subset of the Fellowship objectives. As well as forming a professional qualification in its own right, Associate Fellowship forms an intermediate qualification to full Fellowship.

Again following the development of the Teacher Accreditation Scheme, an Accreditation Scheme is being developed for support and allied staff. Given the wide variation in the roles of such staff, this Scheme has taken a rather different approach. It does not attempt to define all the competences and values of all the jobs in a university. Rather it tries to define what is distinctive about doing these jobs in a university rather than in another type of organization. The Scheme is currently being piloted in six universities, with administrators, technicians, support staff and managers. A December 1996 launch is planned. By focusing on what one might call the particular university-ness of the work of staff, this Scheme may go some way to meeting in...
an appropriate way the view which Brew and Boud (1996) express that 'further emphasis on the specific learning outcomes from [programmes of staff development for allied staff]... is needed if they are to be recognized as part of an educational program for participants.'

From teacher to developer

The SEDA Fellowship Scheme identifies staff and educational developers as having different skills from teachers in higher education, although there is considerable overlap: in terms of skills; to a great extent in terms of values and principles; and also in that most staff and educational developers were once, and in some cases still are, teachers of students. However, a spectrum can be identified from teacher to developer. Teachers become innovators in their own practice, then start to work with colleagues to extend innovations. John Cowan, Director of the UK Open University in Scotland and External Examiner to the SEDA Fellowships Scheme, in conversation with us has used the term 'educational developers' to describe such educational innovators, and has helped us to develop this idea of a bridge from teacher to (as described for the Fellowship Scheme) educational developer. SEDA’s Accreditation Schemes do not yet provide a comfortable route over this bridge. One way forward may be an advanced teacher accreditation scheme, perhaps along the lines suggested by Keessen, Wubbels, Van Tartwijk and Bouhuijs (1996).

The established lecturer

'Teacher as developer', as suggested above, may be one way in which the Teacher Accreditation Scheme can attempt to characterize the rich and subtle range of skills, attitudes, values and principles which experienced and able university teachers bring to their work. Other approaches are likely to be necessary. Jenkins (1996) describes the strong discipline basis which lecturers bring to their teaching – it may be necessary to develop discipline-specific accounts of the competences of an established teacher. Established teachers are also often course leaders and course designers. Are these core skills for an experienced university teacher? SEDA has not yet taken this development very far. One possibility is a modular framework, in which an individual cluster of subject-specific teaching skills and methods, and teaching responsibilities such as course leader, is accredited separately for each teacher.

A role in the NVQ process?

National Vocational Qualifications are currently being contemplated for UK universities (Ecclestone, 1995). The first stage is the development of what is called an occupational map for UK higher education. This will describe the key purpose of the industry concerned (in this case, higher education) and the main functions required to achieve this goal. The development of such an occupational map does not guarantee that NVQs will subsequently be established, and UK universities and their representatives have not expressed universal enthusiasm for the move. The contract for this occupational mapping has now been let, and SEDA is being consulted in the drafting of the map.

A version of the Scheme for librarians?...

Librarians have always had a role in helping students to learn. A UK project to develop the electronic library (eLib, current) has, as one component, the development of a training programme for librarians who train library users in the use of IT in libraries. Rather than simply develop a training programme, this EduLib project has asked SEDA to devise an accreditation scheme and a series of training workshops and associated materials. The accreditation scheme will be very similar to that for university teachers. The University of Hull is one of two lead universities in the EduLib project, and the EduLib programme for librarians at Hull may be a negotiated variant of the SEDA-recognized programme for teachers at Hull.

...for learning technologists?...

A major investment has been made in the UK (as in many other countries) in the development of technology-based teaching and learning materials (see, for example, TLTP, current). The development
of such materials involves a range of skills on the part of the learning technologist, embracing both technological and pedagogic skills. SEDA is exploring with the Association for Learning Technology (see ALT, current) the development of an accreditation scheme for learning technologists. Two variants are being considered—one based on the Teacher Accreditation Scheme, with additional competences, aimed principally at the learning technologists who use, and perhaps also develop, materials in their own teaching; and one based on the Fellowship Scheme, for those who help other academics make use of learning technology.

...for research students who teach?...for part-time teachers?

Teaching assistants and research students provide an increasing part of the teaching for undergraduate students in higher education worldwide. Such teaching assistants (TAs) can bring considerable life and vigour to their teaching, and an empathy with their students and with the nature and sources of student incomprehension which more experienced lecturers may struggle to achieve. Many TAs see their teaching as a step on the road towards becoming a full-time lecturer.

For almost all TAs, the early teaching sessions are, understandably, scary. Many universities are realizing that the teaching undertaken by such students can pose quality assurance problems similar to those posed by full-time lecturers. Great effort is being applied, particularly in North America, to training Teaching Assistants (see, for example, Gokcora 1996, Lewis 1996, and The Journal of Graduate Teaching Assistant Development). Much the same is true for part-time teachers, who are in some ways even less visible and even less supported. SEDA is exploring an approach to accrediting such teachers, again through recognizing training courses, again providing a self-contained professional qualification, as well as a step towards full SEDA accreditation.

Specialist schemes

The objectives and values of the Teacher Accreditation Scheme are designed to be generic, to apply to anyone who teaches in higher education. However, teachers in higher education are almost always subject specialists. Specialist programmes are being accredited. The SEDA-recognized programme at the Bartlett School of Architecture at University College London trains and accredits teachers of architecture. A second programme at the University of Brighton will accredit health teachers and social work professionals. A specialist scheme is being developed by a professional association for teachers of psychology. A programme at the Open University in Scotland accredits distance learning tutors. Other specialist schemes may be developed with the appropriate professional and subject bodies.

Conclusion and discussion

The SEDA Teacher Accreditation Scheme must, on a numerical criterion at least, be judged successful. Some 60 higher education institutions have signed up for this entirely voluntary scheme; old universities, new universities, specialist colleges and, more recently, a polytechnic in Singapore and two universities in Australia. A recent review of the Scheme shows it to be working well, albeit with the need for some operational clarifications and changes. Why has it attracted such support?

Two sets of reasons can be identified—reasons from the external world of higher education, and reasons from within the Scheme itself:

The main external reason is the existence of pressures to increase the quality of teaching and learning in UK higher education, or at any rate to increase accountability for quality. Quality Audit involves study of institutions' own quality assurance systems, to ensure that these are working well and to make suggestions for improvement. Quality Assessment involves study, on a subject-by-subject basis, of courses and how they are taught—classes are inspected by subject teachers. The two systems will be combined under a new quality agency, but early indications are that the new combined system may include many of the features of both current systems.

Both Auditors and Assessors take an interest in teaching quality, and in the means taken by institutions to ensure and improve teaching quality. The existence within an institution of a nationally recognized programme for the initial training of teachers is generally thought to aid a
good Audit or Assessment outcome, as well as having good effects on the quality of teaching itself.

A number of internal reasons for the success of the Scheme can be adduced from informal conversations and from written accounts by programme leaders:

The Scheme’s explicit base in values as well as competences or objectives finds wide favour.

The Scheme is flexible, and can be tailored to the needs and the culture of the institution. Recognized programmes vary very widely indeed; in duration (currently from one semester to three years); in style (from a conventional taught course to individualized tutoring); in arrangements for validation (from conventional university validation through joint SEDA/university validation to recognition only by SEDA); and also in academic level. SEDA does not specify an academic level for SEDA-recognized programmes, describing SEDA accreditation as a professional rather than an academic qualification. In practice, most institutions have rated their programmes as M (Master’s) level.

The Scheme includes a process of staff development for programme leaders. When an institution registers its programme for new teaching staff with the Scheme, two members of the programme team receive a day’s training. During this they work with other programme leaders and course team members, comparing notes on their programmes and developing their understanding of the SEDA requirements and recognition process. The training events also train programme leaders to act as recognizers for other programmes. As soon as their programme is recognized, they in turn recognize others’ programmes. The Scheme is thus a true peer-recognition process.

The recognition process is developmental, rather than confrontational; recognizers themselves run recognized programmes, and the Scheme does not build a cadre of detached ‘expert’ recognizers.

The programme is affordable – the current fee for recognition is £1,500.

What of the future?

SEDA has watched with increasing surprise the growth of the Scheme. The organization sets modest targets each year, only to see them exceeded. The Scheme was originally developed in the London and South East region of the UK – it has spread much wider. It was assumed that virtually all the universities signing up would be new universities – a steadily increasing number of older universities has joined. And there was no thought at all of the programme being adopted outside the UK.

Possible threats include the development of a National Vocational Qualification (NVQ) for teaching in higher education. The higher education sector has thus far shown little enthusiasm for this, although the weight of the NVQ ‘machine’ must not be underestimated. However, even if an NVQ is developed, it would be difficult to see how it could not be influenced very substantially by the current SEDA programme.

The lecturer’s union, the Association of University Teachers, has recently explored accreditation with its members, through a discussion paper (AUT 1996) and a conference – at which SEDA’s Teacher Accreditation Committee Chair, Liz Beaty, described the Scheme (Beaty, 1996). The idea of accreditation is still under discussion by AUT; the significance of this is that the idea would surely have been rejected only a few years ago.

One-quarter to one-third of British higher education institutions (depending how they are defined) have signed on for the SEDA Teacher Accreditation Scheme. Applications and enquiries continue at a steady rate.

Finally, what lessons have we have learned from this development?

Perhaps that:

- Defining the attributes of a higher education profession such as teaching, in terms of values and principles as well as objectives and skills, is found widely acceptable;
- Broad definitions of teaching which allow, and indeed require, interpretation to suit local circumstances, can prove acceptable in a way that tighter definitions may not;
- It is possible to find wide agreement; first among staff and educational developers, then among their institutions and also, most important, among new lecturers in higher education; about the core of the profession or the role of a teacher in higher education.
References


cLib (current). World Wide Web source at http://ukoln.bath.ac.uk/elib/


SEDA (current). World Wide Web source at http://www.seda.demon.co.uk


SEDA (1996a). The Accreditation of Teachers in Higher Education. Birmingham: SEDA.

SEDA (1996b). SEDA Fellowship – Professional accreditation for staff and educational developers in higher education Birmingham: SEDA.


The authors

Carole Baume is Principal Lecturer in Educational Development and Development Officer in the Centre for Staff and Learning Development at Oxford Brookes University. She is also Co-Chair of the Staff and Educational Development Association and a founding member of SEDA’s Teacher Accreditation Committee. Carole edited the first Teacher Accreditation Yearbook.

David Baume was until recently head of the Educational Development and Support Service at London Guildhall University. From February 1996 he will be Co-Director of the new Centre for Higher Education Practice at the Open University. He is a former Chair of the Staff and Educational Development Association.

Carole and David have acted as external validators and examiners for many programmes for new teaching staff in UK Universities. They have recently written a series of booklets for research students who teach, published by the Oxford Centre for Staff Development.

Address: 15 Hanmer Road, Simpson, Milton Keynes, MK6 3AY, England
Tel: +44 (0)1908 672 830; Fax: +44 (0)1908 674 954
E-mail: 100257.70@compuserve.com

A summary of a detailed evaluation of a one-term development project at Oxford Brookes University intended to explore, and where appropriate to extend, the use of information and communications technology for teaching and learning.
IT Term: a model for institutional change?

Carole Baume & Alan Jenkins, Oxford Brookes University and David Baume, The Open University

Achieving institutional change

Achieving CTI's mission to maintain and enhance the quality of learning... through the application of appropriate learning technologies requires activity in many areas. In the UK and elsewhere, much of that activity has focused (through TLTP and CTI) on the development of IT-based materials, and on their take-up by enthusiasts in the disciplines. Important as these activities have been and are, there is increased recognition that they need to be complemented by strategies that focus on institutional change. Thus the recent Coopers and Lybrand et al. evaluation of TLTP (see page 68 for Executive Summary) emphasised the critical importance of institutional change to achieve the dissemination and integration of courseware and other IT-based initiatives into student learning.

To promote such institutional change, Oxford Brookes designated the summer term of 1996 as IT Term. This involved a considerable range and number of university wide and School/ Departmental based events at which staff could see, talk about and experiment with IT-based applications to teaching and learning. We consider this idea to be one model for institutional change that could be adapted by other institutions. Before suggesting its implications for other institutions, we need to briefly set out the model and then describe what happened during IT Term. In reading this account, you need to know that two of us - Carole Baume and Alan Jenkins - were in the University's staff development unit and were responsible for much of the University wide events and policy. David Baume was hired by TLTP and Oxford Brookes to evaluate the term independently.

The model summarised

The central features of the idea/model were to:

• Dedicate an extended period of time - in our case an eleven week term - to an institution-wide programme of activities focussing on IT. The original idea was to hold one event for perhaps a day or week. We decided that to achieve significant institution wide change one required a much more extensive period of time. We decided a term gave enough time, but also ensured an intellectual and organisational intensity which, without significant extra resources, might not be able to be maintained over a longer period.

• Decide that these activities are to be staff development events concerning IT and student learning - ie we did not involve students in these events. We wanted staff to try out, to debate and to explore educational uses of IT in some privacy and without any requirement that they adopt these methods. We also wanted to ensure that the focus was on the staff development of the many, rather than on the purchase or development of courseware for or by a few staff.

• Ensure a wide range of university-wide and school/departmental events - we felt this to be appropriate to Brookes' institutional culture. This culture is one in which, partly through an institution-wide modular course, there is a strong sense of being part of one institution. However, it is an institution in which subject groups and schools enjoy considerable autonomy and are very resistant to institution-wide managerial solutions. Clearly other institutions adapting IT Term will need to adapt the idea to their institutional culture/organisation.
Commit top-sliced institutional funds to support the term. Oxford Brookes allocated £25,000 to support the term. Most of this money was allocated to the 16 academic Schools (eg Social Sciences) and selected Support Departments (eg the Library). Each was allocated £1,000. However, they only received that money when a centrally appointed group agreed that their plans for the term met basic criteria, ie being about staff development for the many not the select few, and focused on IT and student learning. None of the centrally distributed money could be used for buying equipment, software etc.

What happened in the term
The term featured three main types of activities:

- university wide spectaculars;
- university wide theme events;
- school/department activities.

The spectaculars included:

- A two day opening event organised in conjunction with the CTI Support Service, which brought all CTI Centres to Brookes. This clearly demonstrated the range of materials now available, and opening talks by Jonathan Darby and Wendy Hall highlighted key issues in the integration of IT.
- A Hypothetical exercise led by the university’s Chancellor, Helena Kennedy QC. This evening event borrowed the television format of getting professionals to state how they would react to a hypothetical scenario. This scenario centered around a student entering Brookes in 2001, involved senior management and senior staff from outside Brookes (including Sarah Turpin of TLTP and Robin Middlehurst of HEQC) publicly debating the future role of IT at Brookes. A video of this event and support materials is available for purchase.
- The third spectacular involved the TLTP support network organising with Brookes a day which involved a mixture of presentations and, perhaps more critically, meetings of senior committees at Brookes with the TLTP support network on the policy issues involved in the integration of IT into student learning and institutional cultures. The other main institution wide set of activities was low key and small scale.

Throughout the term there was a mixture of open seminars and workshops on aspects of IT that we considered particularly relevant to immediate issues at Brookes. These issues included the roles and implications of IT in equal opportunities, assessment, modular administration and student guidance.

The third set of events, and perhaps the most fundamental, were the locally-organised events in schools and support departments. The activities were very varied and addressed issues of immediate concern to discipline-based staff. The School of Biological and Molecular Sciences had an away day to review how IT skills were developed and supported in their courses. Humanities developed an annotated guide to the use of the Internet for staff and students in History, English and Art History. Health Care Studies activities included a video conference discussion on curriculum issues with staff in Colorado.

The evaluation indicates the value to Brookes of the three-pronged approach in effecting change. The spectaculars gave a clear signal of the importance of the issues and made everyone aware of IT Term. The willingness of senior management to commit resources and engage in public debates such as the Hypothetical was recognised and valued. The locally organised events in schools are now impacting on school-based strategic planning. In short we are certain that IT Term was a valuable activity for Brookes. What are our suggestions for other institutions who wish to adapt this idea?

Suggestions for elsewhere: adapt it for your institution
Clearly you need to adapt our model to your institution. Amongst the issues to consider/recognise are:

What is the appropriate timing of the event?
Brookes decided on a term, because, in a three term calendar, a term was long enough to mount a range of activities, but also gave us a clear focus. However, by the end of the term we were suffering IT Term fatigue.
You certainly need to recognise that much time is needed to plan the event. At Brookes there was a nine-month lead in to the event after the formal agreement by Academic Board. You also need to consider what to do after the term, to ensure that the impetus is not lost.

Recognise that this will take up many resources, including staff time.

Though clearly there is no need, and perhaps also no opportunity, for other institutions to have such a large number of events, you need to recognise (we wish we had) the considerable time and effort it takes to organise just one event. The key people whose time has to be recognised and valued include the administrators at university and school level, and staff in computer services who will take on much of the detailed organisation and must receive the rewards of public recognition.

Decide what is an appropriate mixture of university-wide/school/department based activities.

Your institutional culture and organisation will shape what is the appropriate (and possible) mixture of university-and-school based events. We suggest using as a wide a variety of appropriate methods as is possible.

Opportunistically base your plans upon your institutions structures, strengths and interests.

At Brookes the idea grew out of our particular committee structure: the Teaching and Learning Technology Group - that united the different concerns of the Computer Centre, Library, staff development and school-based staff. We then ensured that it was taken up by various other committees and in particular that it was owned by the devolved institutional culture and organisation.

We opportunistically built one spectacular, the Hypothetical, around Helena Kennedy QC, our Chancellor. Opportunistically, we also linked the term to an institution-wide commitment to equal opportunities and profiling. Other institutions have their own particular organisational features and strengths, eg close links with local/national based IT companies, an interest in progressing their research culture or a commitment to changing assessment practices. In adopting the idea you need to adapt it to build upon your local features.

Include elements of debate and watch the technological hype.

Perhaps this suggestion reflects particular features at Oxford Brookes, including the limited areas of the University that are technology based and that one of the organisers (Alan Jenkins) was and remains a Luddite? However, we do suggest that the adoption of IT has been hindered by the public emphasis on technological efficiency, which is simply not yet demonstrated by results. We also suggest that in a university there should be a proper emphasis on discussion and debate and that institutional change to the adoption of IT needs to build in space and opportunities for staff to explore and challenge the relevance of IT. At Brookes we had the opportunity of a Chancellor who could dramatise this issue. Were we taking the idea elsewhere we would look for local opportunities and structures to build in such a debate.

Think through how to follow up the term.

As we developed the idea and structure for the term we became concerned to ensure it was not a one-off event.

The strength of the model lies partly in its dramatic nature and in it being a focus and fulcrum for change. However, there is the linked danger of it being but another institutional initiative which is soon forgotten and replaced by some other concern. You need to ensure that its structure includes elements that encourage or require groups to follow up what they learned from the term. One way we attempted to do this at Brookes was to require all Schools and Departments to identify actions they intend to take after IT Term, and how they consider university policies could best support these initiatives. We are now working to progress these ideas to better ensure that IT Term has contributed to an institutional change in the application of appropriate learning technologies.

Further information on IT Term

If you think this account has indicated that this model can be adapted to your institution, you can obtain full details on the terms activities and a copy of the evaluation report through our special IT Term Web site at http://www.brookes.ac.uk/IT-term/

Copies of the video of the Hypothetical including supporting material are available from Marj Bolton, OCSDL, Oxford Brookes, Oxford, OX3 OBP, price £15 including postage and packing. Cheques should be made payable to Oxford Brookes University.

References


Publication 3  

A detailed guide for participants on an Open University course in teaching in higher education. The guide explains in detail the course outcomes, learning process and assessment. It also illustrates how the course outcomes may be achieved in relation to a participant's own teaching.
The Portfolio

Portfolio: Guide
Building your portfolio
Presenting your portfolio
Portfolio: Guide

Prepared for the course team by David Baume
The H851/HH851 course team

David Baume (CeHEP) (production chair)
Carole Baume (CeHEP) (presentation chair)
Fiona Carey (Publishing) (editor)
Graham Gibbs (CeHEP) (Reader and Practice Guides academic editor and author)
Sue Glover (editor)
Chris Gravell (Publishing) (editor)
Barbara Hodgson (IET) (programme director)
Agnes Kukulska-Hulme (IET) (FirstClass online conferencing co-ordinator)
Anne Langley (Student Services: Planning) (reader)
Ranald Macdonald (CeHEP) (tutor recruitment co-ordinator and reader)
Mary Masterton (IET) (course manager)
Diane Mole (Graphic Design) (designer)
Andy Northedge (IET) (author)
Alistair Morgan (HSW) (author)
Clive Pearson (CeHEP) (website materials co-ordinator)
Derek Rowntree (IET) (reader)
Howard Senter (editor)
Mary Shepherd (OU associate lecturer) (reader)
Sue Roberts (IET) (course team secretary)

The course team gratefully acknowledges the participation and support of Helen Boyce, Karen Cvijetic, Joanne Deane, Amanda Gilbert, Robin Mason, Brenda Parish, William Prescott, Margaret Robertson, Alison Whittaker and John Willmer.

The Open University, Walton Hall, Milton Keynes MK7 6AA
First published 1998
Copyright © 1998 The Open University
All rights reserved. No part of this publication may be reproduced, stored in a retrieval system or transmitted, in any form or by any means – electronic, mechanical, photocopying, recording, or otherwise – without either the prior written permission of the publisher or a licence permitting restricted copying issued by the Copyright Licensing Agency, 90 Tottenham Court Road, London W1P 0LP. This book may not be lent, resold, hired out or otherwise disposed of by way of trade in any form of binding or cover other than that in which it is published, without the prior consent of the publisher.

Edited, designed and typeset by The Open University.
Printed in the United Kingdom by The Open University
SUP 38451 9

This text forms part of an Open University course. If you have not enrolled on the course and would like to buy this or other Open University material, please write to Open University Educational Enterprises Ltd, Offices 7, The Open University, Walton Hall, Milton Keynes. MK7 6AA, United Kingdom. If you wish to
How to use this Portfolio:Guide

Live with it. This is not a Reader – it is a working document to which you will constantly refer.

You may find it helpful to glance through the Portfolio:Guide to help you locate particular sections. It is a reference book. You will find you need to keep returning to it to help you to analyse your evidence and decide how to show that your work meets the criteria for accreditation.

It is a complicated document at first sight. But, if you use it alongside the process of collecting and analysing evidence, it will provide a valuable support for the development of your portfolio.

Mary Shepard

Mary is a social sciences lecturer who has used an early draft of this Portfolio:Guide to put together a portfolio.

Author’s acknowledgements

H851, Teaching in Higher Education, is the first course from the University’s new Centre for Higher Education Practice. Chairing the production of H851, Teaching in Higher Education, has been my first experience of an Open University course team. Staff of the Centre all had a lot to learn, and learn very quickly. We could not have done this without the support of a large number of very able and helpful people.

The academic and pedagogic attention applied to this Portfolio:Guide was much greater than I have ever received before in developing teaching materials. Whatever its current merits and defects, the Portfolio:Guide is much better for the many hours of constructive and critical scrutiny which it received from the course team.

Graduate teaching assistants and lecturers from Sussex, Lincolnshire and Humberside and elsewhere tested the Portfolio:Guide (and other course materials). Their sometimes robust but always supportive critiques have had a very positive effect on the materials. My thanks to: Maaruf Ali, Julian Beckton, Jo Bell, Helen Bowman, Louise Bywood, Kate Cavanagh, Stephan Collishaw, Paula Cooper, Richard Darlington, Kris Fletcher, Brian Garrigan, Natalie Haigh, Jon Hanna, Ashleigh Hillier, Steve Hudson, Shorful Islam, Suzanne Kear, Jenny McConnell, Kate Parkinson, Joanne Saddlington, Clara Strauss.

The administrative and support staff of the Centre for Higher Education Practice contributed considerably beyond the call of duty.

Our course manager and lead editor have provided us with superb support.

I am sure that the Portfolio:Guide can be considerably improved for further presentations of the course, and that the Portfolio:Guide model can be further developed for future courses. I thank you, as a participant in the course, in advance, for telling us how we can make these improvements!

David Baume
Course team chair (production)
a.d.baume@open.ac.uk
## Contents

1 Introduction

1.1 Welcome

1.2 What will the PortfolioGuide do for you?

1.3 How will you use the PortfolioGuide?

1.4 Frequently asked questions about building a portfolio

1.5 A summary of the assessment process

2 The bases on which you will be assessed

2.1 Outcomes

2.2 Values

3 Evidence about your teaching

3.1 Validity

3.2 Information you may receive or obtain

3.3 Information you may generate

3.4 Simulated evidence

3.5 Specifying, collecting and sorting evidence, and a form for logging evidence

3.6 Showing principles and values in action

3.7 Judging evidence

4 Making a claim for accreditation

4.1 Outcome 1 – Designing teaching sessions

4.2 Outcome 2 – Using teaching and learning methods

4.3 Outcome 3 – Using assessment methods

4.4 Outcome 4 – Monitoring your work

4.5 Outcome 5 – Keeping appropriate records

4.6 Outcome 6 – Using coping strategies

4.7 Outcome 7 – Reflecting and planning CPD

4.9 The overall assessment of your portfolio

4.10 What to do next?

5 Submitting your portfolio for assessment, and what happens next

5.1 Administrative arrangements

5.2 Further studies in higher education practice

6 Further information

6.1 Publications and other sources

6.2 The accreditation of teachers in higher education
1 Introduction

1.1 Welcome

Welcome to the Portfolio:Guide for H851, Teaching in Higher Education, and for the APEL version of this course, HH851. This Portfolio:Guide takes you through the process of demonstrating your competence in teaching in higher education.

Our aim is to help you to produce a successful portfolio so that you pass the course and gain accreditation. The process of completing this portfolio should be valuable in its own right. So should the portfolio itself. The portfolio will provide evidence of your teaching ability which you could use for appointment and promotion, and give a basis against which you can measure your continuing professional development as a teacher. We hope it will become an object of quiet professional and personal pride.

What do we mean by a portfolio? We mean an organized and annotated collection of documents, such as module outlines, lesson plans, feedback to students, diary pages and work records and notes from colleagues. We call such material 'evidence'. This evidence backs up the series of reflective pieces which explain why you teach in the way you do. We call these your 'claim'.

Underpinning this course is the idea that students deserve good teaching. Teaching well can be a delight, as well as being intellectually and personally challenging. Teaching less well, or simply not knowing how well you are teaching, can be dispiriting. The process on which you have embarked will help you to know whether you are teaching well, and whether you are continuing to improve.

Course outcomes and values

What do we mean by a good teacher? We mean someone who has shown that they can:

1. design teaching sessions;
2. use appropriate teaching and learning methods;
3. mark or grade, and give feedback on, student work;
4. monitor their own teaching;
5. keep appropriate records of their teaching support and academic administration;
6. manage their time, and operate successfully within available resources;
7. reflect on their work and plan their continuing professional development.

These are the outcomes of this course.

We also mean someone who can do all this in a way that is clearly informed by:

1. an understanding of how students learn;
2. a concern for students' development;
3. a commitment to scholarship;
4. a commitment to work with and learn from colleagues;
5. the practising of equal opportunities;
6. continuing reflection on professional practice.

These are the principles and values which must underpin attainment of the outcomes. For example, the teaching sessions which you plan must help students to learn and to develop. And you must monitor your teaching in a scholarly way, making use of the results of research into obtaining and using feedback from students. You will thus demonstrate not just that you can teach, but that you can, in defined ways, teach well.

Successful completion of this course will gain you 30 M-level credits, and professional accreditation by the Staff and Educational Development Association as an Associate...
Figure 1 Assembling your portfolio, from evidence collection to claim.
1.2 What will the Portfolio:Guide do for you?

This Portfolio:Guide has been prepared for those who are undertaking the complete H851 course, and for those who seek to have their previous learning accredited through HH851. The assessment process in each case is the same.

Within this guide, each section does the following:

- Section 2 describes the learning outcomes and the underpinning principles and values which you will be required to demonstrate;
- Section 3 guides you through the process of finding, making, collecting and collating evidence of your teaching;
- Section 4 helps you to use this evidence to make a claim for assessment and to evaluate the claim you have made, suggesting what to do next if you feel that your claim doesn't yet stand up;
- Section 5 tells you how to submit your portfolio for assessment;
- Section 6 provides further useful information and sources.

At the end of this guide is the start of your own area of the ring binder, where you assemble your portfolio. There are two sections in this area: 'Building your portfolio' contains a bank of blank forms to help you collect, collate and analyse evidence and write and test your claim for accreditation; 'Presenting your portfolio' contains the Portfolio Assessment Form and is where you should insert, now, the numbered dividers you received with this guide, and in due course your completed portfolio for assessment.

1.3 How will you use the Portfolio:Guide?

In your work as a teacher, you will both receive and generate information about your work.

Information you receive will include:
- course syllabuses;
- work and feedback from students;
- notes of observations of your class by a peer, mentor or manager.

Information you generate may include:
- lesson plans;
- lecture notes;
- visual materials such as OHP slides;
- handouts;
- feedback to students;
- e-mail messages;
- marked student work.

Some of this information shows you working on and developing your skills as a teacher.

This Portfolio:Guide will help you to collect, and then to select and sort, appropriate evidence. When you make your claim for your successful completion of the course, and thus for accreditation as an associate teacher, this portfolio will contain the evidence to back up your claim.

Making the claim for accreditation involves more than collecting evidence. It also involves you in analysing this evidence. This analysis will be much easier if the evidence is already filed under useful headings. The course learning outcomes and the underpinning principles and values define the portfolio headings we have adopted here.

Why have we called this a 'Portfolio:Guide'?
Introduction

In more detail, it starts its life as a guide on how to collect and sort evidence for your portfolio. It also provides a receptacle to put this evidence as you collect it, and it suggests what kind of evidence goes where. It provides examples of evidence, and examples of claims that the evidence shows that the outcome has been met.

1.4 Frequently asked questions about building a portfolio

Q How much evidence should I include?
A The maximum number of pages of evidence allowed for each outcome is specified on the Portfolio Assessment Form. (The maximum total is 180 pages.) I suggest that you collect more than this, and then sift and select as you write your claim.

Q How long a claim should I write?
A Again, the word limits are specified for each outcome on the Portfolio Assessment Form. They total some 5000 words.

Q Some of my ‘evidence’ is just scribbles in a notepad, and some of it isn’t written down at all – for example, I have just memories of important conversations. What should I do?
A Write up the scribbles or the conversation, stating that you have written them up subsequently. Say when the original notes were made or when the conversation happened. When you do this, say what is a record and what is later interpretation.

Q Some evidence isn’t very clear as it stands – how do I show where it came from or how I used it?
A Introduce each piece of evidence. Say, for example, what it is; why you produced it; when; and in which course, with which students, how and in what setting, you used it. Include this material with the evidence, not the claim.

Q What about work I do with other teachers? Can I include this?
A Absolutely! Collaboration is one of the underpinning values of the course. Just add a note to each piece of work that you did with someone else, explaining what you did and what they did. Sometimes you may need to add a note from your colleague confirming who did what.

Q Can I include evidence from my teaching work in more than one institution?
A Yes indeed. You are working towards a national professional qualification in teaching in higher education. It is fine to include evidence gained in more than one institution.

Q I only teach laboratory classes. Is this enough?
A I have to answer this with a question – does your teaching work enable you to do each of the things described in the course outcomes in Section 2? Do you: design teaching sessions (in this case lab sessions); use two appropriate teaching and learning methods (in a lab these may for example be making short presentations as well as facilitating the laboratory classes); and mark or grade and give feedback on student work? (The last four outcomes listed in Section 2 can be attained in any teaching setting.)

Q I tend to plan a class, run the class, get feedback, and then change the plan for next time. How do I show this in my portfolio?
A Put together your original plan, your notes and feedback, and the revised plan, and some notes on why you made the changes you did. Keep these together.

Q Can I use some evidence to show more than one outcome attained?
A Certainly. Copy the evidence into the section for each outcome to which it refers. The page limits for each outcome still apply!

Q How long will it take me to put my portfolio together?
A People who have put together pilot portfolios using this Portfolio:Guide tell us
1.5 A summary of the assessment process

<table>
<thead>
<tr>
<th>Activity</th>
<th>Portfolio/Guide section</th>
<th>Other resources you will need</th>
</tr>
</thead>
<tbody>
<tr>
<td>To be completed by...</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 List your main teaching activities. These may be what you have done in the past two years; or what you are currently doing; or what you expect to undertake during the period in which you are preparing for assessment for H851/HH851.</td>
<td>Section 2 describes outcomes and values.</td>
<td>Wall chart shows outcomes and values.</td>
</tr>
<tr>
<td>2 Check that this work should enable you to demonstrate each of the outcomes and their underpinning values, and that you will be able to generate and accumulate the necessary evidence.</td>
<td>Section 3 describes evidence.</td>
<td>Mentor or colleague(s).</td>
</tr>
<tr>
<td>If your teaching experience or your planned teaching work doesn't enable this, you face two main alternatives. You can:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A Work with your mentor, or negotiate with a colleague, to undertake additional or different teaching to enable you to achieve these outcomes, or</td>
<td></td>
<td></td>
</tr>
<tr>
<td>B Plan how you will demonstrate part of an outcome by simulation, subject to the restrictions on simulated evidence described in Section 3.4.</td>
<td>Section 3.4</td>
<td></td>
</tr>
<tr>
<td>3 Plan what evidence you should be able to collect about your teaching that should help you show you have attained each of the outcomes.</td>
<td></td>
<td>Preliminary Evidence Planner in 'Building your portfolio'</td>
</tr>
<tr>
<td>4 Collect the evidence and file it in your portfolio under the appropriate learning outcome. We have provided a form which you can use to log evidence and make an initial judgement of which outcomes and which values it shows you attaining.</td>
<td></td>
<td>Evidence Organizer in 'Building your portfolio'</td>
</tr>
<tr>
<td>5 Analyse the evidence to see how well it shows that you have achieved each outcome in a way that is underpinned by the appropriate values.</td>
<td></td>
<td>Section 4</td>
</tr>
<tr>
<td>6 Assemble the evidence for each outcome. Further forms to help you do this are also provided.</td>
<td></td>
<td>Outcome Evidence Organizers in 'Building your portfolio'</td>
</tr>
<tr>
<td>7 Write a claim or self-assessment for each outcome. In this self-assessment you need to say how the evidence you have collected demonstrates that you have attained the learning outcome. You also need to say how the evidence shows the appropriate values in action.</td>
<td></td>
<td>Colleagues/mentor</td>
</tr>
<tr>
<td>8 Assess the claim and the evidence, alone or preferably with advice from colleagues or a mentor, to see whether the evidence and claim make the case convincingly. The Assessor Guide includes notes on how to assess – you should follow this guidance and assess your own portfolio to check whether you need to do further work.</td>
<td></td>
<td>Portfolio Assessment Form in 'Building your portfolio'</td>
</tr>
<tr>
<td>9 Depending on the results of your self-assessment, you should follow one of three routes:</td>
<td></td>
<td>Confirmation of Authorship and Outcome Evidence Records in 'Presenting your portfolio'</td>
</tr>
<tr>
<td>A If your review was positive, submit your portfolio at the first assessment date.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>B If some of the evidence and/or your teaching was weak, you may decide that you need to choose the next assessment date to allow you to develop your teaching further and/or collect more evidence.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C If you did not find evidence for some outcomes, you will need to be assessed later to enable you to undertake the necessary teaching work.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10 When you have submitted your portfolio for assessment, your assessors will decide whether you have successfully completed the course, or whether you should undertake further work. If you are asked to undertake further work, you will be told which outcomes and/or professional values you have not yet demonstrated.</td>
<td></td>
<td>Notify the OU by 23 March 1999 if you do not wish to be assessed in May 1999.</td>
</tr>
</tbody>
</table>
2 The bases on which you will be assessed

You will be assessed on two different but closely related bases:

- a set of learning outcomes;
- a set of underpinning principles and professional values.

These learning outcomes and values form part of the accreditation requirements of the Staff and Educational Development Association (SEDA), the body that will accredit you on successful completion of this course. More information on teacher accreditation is given in Section 6. The outcomes are similar to those being developed for the proposed Institute for Learning and Teaching.

2.1 Outcomes

You will need to show that you can carry out the following activities in order to complete the assessment successfully, and thus be accredited.

These learning outcomes are that you should have:

1. Designed teaching sessions from a course outline, document or syllabus. This involves choosing teaching methods appropriate to the group of learners, the mode of study, the subject material, the resources available and the learning outcomes.

2. Used two of the following appropriate teaching and learning methods: making presentations (e.g. lectures, demonstrations); facilitating group learning (e.g. through seminars, discussion groups, projects); working with individual learners; facilitating practicals or laboratory classes. You must also include evidence of having used appropriate learning technologies.

3. Marked or graded, and given feedback on, student work.

4. Monitored and evaluated your own teaching, using own, peer and student feedback.

5. Kept appropriate records of your teaching support and academic administrative work.

6. Developed personal and professional coping strategies appropriate to the constraints and opportunities of your institutional setting, to manage your time adequately and operate successfully within available resources.

7. Reflected on your own personal and professional practice and development, assessed your future development needs and made a plan for your continuing professional development (CPD).

2.2 Values

Good teaching requires more than exercising a set of skills. Good teaching is informed by some values or principles.

Why are these necessary?

Consider equality of educational opportunity.

A teacher could show a sound knowledge of their institution's equal opportunities policy. Despite this knowledge they might still teach or assess in a way that denies equality of opportunity to a particular individual or group of students in their class.

It is thus not enough to know about equal opportunities. It is necessary that concern for equality of opportunity affects the teacher's decisions and actions.

The same is true for a concern for students' learning and development; for taking a scholarly approach to teaching; for collaboration with colleagues; for continued critical
The following values (quoted from the SEDA Teacher Accreditation Scheme) must demonstrably underpin your practice as a teacher.

1 An understanding of how students learn

All teaching and academic administration should be informed by an understanding of how students learn and the conditions and processes that support student learning.

[To be demonstrated as underpinning Outcomes 1, 2, 3, 4 and 7.]

2 A concern for students’ development

Helping students to learn must begin with a recognition that all students have their own individual learning needs and bring their own knowledge and resources to the learning process. Work with students should empower them and enable them to develop greater capability and competence in their personal and professional lives.

[To be demonstrated as underpinning Outcomes 1, 2, 3, 4 and 7.]

3 A commitment to scholarship

At the base of professional teaching is an awareness and acknowledgement of the ideas and theories of others. All teaching should be underpinned by a searching out of new knowledge – both about the subject/discipline and about good teaching and learning practice. All teaching should also lead to students developing a questioning and analytical approach.

[To be demonstrated as underpinning Outcomes 4 and 7.]

4 A commitment to work with and learn from colleagues

Much of an academic's work is carried out as part of a team made up of teaching staff and academic support staff. Collegiality and support of peers is as important as individual academic excellence.

[To be demonstrated as underpinning Outcomes 4, 6 and 7.]

5 The practising of equal opportunities

Teachers must be concerned that students have equal opportunities, irrespective of disabilities, religion, sexual orientation, race or gender. So, everything that teachers do should be informed by equal opportunities legislation, by institutional policy and by knowledge of best practice.

[To be demonstrated as underpinning Outcomes 1, 2, 3, 4 and 7.]

6 Continuing reflection on professional practice

Teachers should reflect on their intentions and actions and on the effects of their actions. They should try to understand the reasons for what they see and for the effects of their actions. They thus continue to develop their understanding and practice and therefore inform their own learning.

[To be demonstrated as underpinning Outcomes 2, 3, 4, 5, 6 and 7.]

Section 4 of this Portfolio: Guide will give you guidance on ways to show that these values underpin your practice.
3 Evidence about your teaching

This section describes some qualities of good evidence and the process of collecting, collating and annotating evidence that you will use to show that you have attained the learning outcomes in a way that is underpinned by the values.

3.1 Validity

Currency

This course leads to accreditation as a teacher; that is, a certification of your current competence as a teacher. Not more than 20 per cent of the evidence should be more than three years old at the date at which you submit for assessment. Each item of evidence should bear the date when it was produced.

Self-assessment

One of the assessment requirements is that you show that some of your work has been subject to self-assessment. In a real sense your whole portfolio is self-assessed. You decide when you have sufficient evidence of the appropriate quality about your work. You also make the claim that this evidence shows that you have attained the course outcomes.

Peer assessment

Another assessment requirement is that you show that some of your work has been subject to peer assessment. You can do this in many ways. You can define your peers to be other participants on H851 or HH851; other teachers in your department or institution; your mentor; or a staff developer. One or more of your chosen peers must give you feedback from their observations of two teaching sessions – you will need this evidence to demonstrate Outcome 2. You must also show how you have used peer feedback to review and improve your teaching – you will use this evidence in the demonstration of Outcome 4.

Additionally, you may use feedback from peers as evidence in any other outcomes, as suggested in Section 3.2 below.

What is it?

Your assessors need to know some things about your evidence to help them judge whether the evidence shows what you claim it shows. What they need to know depends on what the evidence is. However, these things at least should be clear for each item of evidence:

- the course (e.g. 'BA Psychology');
- the subject ('Experimental method');
- the level ('2nd year');
- who produced it ('name');
- how you used it ('I gave it out the week before the seminar');
- what it is ('seminar preparation task').

Also provide any other information you think your assessors will need to help them make sense of the evidence and of the claim. Simply add such information to the evidence – no need for a separate sheet.
3.2 Information you may receive or obtain

At the start of your teaching, to help you plan a course which you are to teach, you may be provided with, or obtain, some or all of:

- A course document or outline, possibly including an account of the course aims and intended objectives or learning outcomes
- A syllabus
- Assessment papers from previous years
- Teaching notes prepared by those who have previously taught the course
- Handouts
- Audio-visual materials
- Lists of references and sources

Before you assess students' work, you may have, in addition to the above:

- Marking schemes
- Assessment questions, tasks and assignments
- Assessment criteria
- Copies of student work with feedback comments

You may also have received:

- University and department policy statements
- Administrative memos and procedures
- Forms

Very little if any of this kind of material will make its way into your final portfolio. However, you should keep and organize such material carefully, as it provides context and background for your teaching work. You may need to quote from it as you describe teaching choices you have made.

Material prepared by others which may make it into your portfolio includes course outlines (from which you prepare teaching sessions), and descriptions of student assignments (from which you may prepare marking schemes).

You should collect feedback on your teaching. Although produced by others, this is uniquely about your work as a teacher. It very properly belongs in your portfolio.

3.3 Information you may generate

You will probably generate a lot of material as you teach. Some of it you will be able to use as evidence. This material may include:

- Lesson plans
- Visual aids
- Handouts
- Worked examples
- Reading lists
- Assessment questions, tasks and assignments
- Samples of marked student work
- Student work on which you have given feedback
- Notes you have made before, during and after a class
- A teaching diary or log
- Feedback you have received from students, peers or your mentor on your teaching
- Your critique, analysis and use of this feedback

3.4 Simulated evidence

Your normal work may not allow you to demonstrate completely some of the outcomes. For example, your work may be confined to just one teaching method. Or
The best solution in such situations is to negotiate with a colleague, the lecturer responsible for a course on which you teach, or your mentor to undertake one or two examples of the missing teaching work for real. (Your offer may well be accepted with enthusiasm!) Failing this, you may be able to negotiate with one of your teaching colleagues to swap some teaching work to your mutual advantage.

Failing all of this, you can offer simulated evidence. By simulated evidence I mean, for example, evidence about teaching work which was planned but not delivered, or assessments of work undertaken by people other than the students for whom you have assessment responsibility. You may not demonstrate any objective wholly by simulation. Simulation is only allowed for parts of Outcomes 1, 2, 3 and 6.

<table>
<thead>
<tr>
<th>Outcomes</th>
<th>Difficulty in providing real evidence</th>
<th>Real evidence for the outcome</th>
<th>Possible simulated evidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Learning outcome 1 - Designing teaching sessions</td>
<td>All I ever get to design is the seminar following the lecture – but I've got some good ideas for how I'd redesign the lecture and seminar slot together.</td>
<td>Your plans for the seminar sessions which you run.</td>
<td>Your plan for the joint lecture/seminar slot which you are unable to implement because you don't have the responsibility.</td>
</tr>
<tr>
<td>Learning outcome 2 - Using two appropriate teaching and learning methods</td>
<td>This year, I'm only demonstrating in lab sessions. I am also giving a series of research seminars to my faculty.</td>
<td>Your reports of lab sessions you have run, with peer observation and student feedback reports.</td>
<td>Your plans, overheads, reports and feedback from your research colleagues on your research seminars.</td>
</tr>
<tr>
<td>Learning outcome 3 - Marking or grading, and giving feedback on, student work</td>
<td>I'm only being asked to give feedback on drafts of student work, not to mark or grade.</td>
<td>Assignment attachment forms, examples of feedback which you have given to particular students with their responses.</td>
<td>Your marking or grading of the work of a colleague's students, and/or your &quot;pretend&quot; gradings of student work on which you gave feedback.</td>
</tr>
<tr>
<td>Learning outcome 4 - Monitoring and evaluating your own teaching</td>
<td>There are no circumstances under which you cannot monitor and evaluate your own teaching!</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Learning outcome 5 - Keeping appropriate records of teaching support and academic administration</td>
<td>Again, you will always be able to keep appropriate records. They may not be extensive, but they should be all of the records appropriate to your teaching situation.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Learning outcome 6 - Developing personal and professional coping strategies</td>
<td>The balance between teaching, research and the rest of my life is causing no problems this year. However, next year...</td>
<td>Your timetable, work plan, extracts from reflective journal for this year.</td>
<td>Your plans for how you will cope with the changed circumstances you anticipate for next year.</td>
</tr>
<tr>
<td>Learning outcome 7 - Reflecting, identifying needs and planning CPD</td>
<td>There are no circumstances under which you cannot reflect, identify your developmental needs and plan your continuing professional development, even if these are not formally required of you by your institution.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The requirements for your simulation are that:

*It should be as close as possible to reality.* For example, if you are simulating the design of a teaching session, then you should really design one, in your discipline, within all the constraints and requirements of teaching this subject in your...
You should prepare all of the associated materials, whether for the teaching session or the assessment, as if it were real.

If you simulate a teaching method you should gather together a group of people – other participants in your institution on this course would be good, this is a service you can do for each other – and run the class. This is a simulation only in that the audience is not the actual intended audience. A research seminar is another possibility. Similarly, for a simulated assessment you need to find some people to undertake the assessment task, you need to mark their work and give them feedback; or you might mark or grade, and give feedback on, student work from a previous year.

3.5 Specifying, collecting and sorting evidence, and a form for logging evidence

At the end of this process of evidence collection you will have a draft portfolio with at least the appropriate amount of convincing evidence in each section.

Unless you are uncommonly well organized – in which case, congratulations, and go straight to Section 3.6 – here are some suggestions on collecting and sorting and making sense of your evidence.

There are two main strategies. The first is to amass material as it comes and organize it later. The second is to make a plan for your portfolio, for how you will show and claim your ability as a teacher under each of the learning outcomes, and then decide what evidence would back up that claim and seek it out. Let’s consider each approach in turn.

Collecting evidence as it comes

You may want to make an extra copy of potentially portfolio-worthy material that you produce and receive.

What you do next is crucial. You can, in the rueful words of someone who piloted this assessment process, ‘sort it all into one large box’. If this is all you do, you will soon have a lot of paper in a large box, several steps away from a portfolio. You should sort your accumulating evidence, probably each week.

What do I mean by ‘sort’?

You need to decide which learning outcome each piece of evidence most obviously shows that you have attained. ‘Building your portfolio’ which follows this guide contains a photocopiable master of the Preliminary Evidence Organizer, a form which you can use to log, sort and file evidence, and start to identify which learning outcome or value it shows in action. Figure 3 shows an example of the form in use.

Planning your collection of evidence

This approach starts, not with evidence, but with the learning outcomes and with the values. You decide what kinds or pieces of evidence would show that you have attained each learning outcome and each value, then you collect or develop such evidence from your teaching.

Or you could start by drafting or outlining your claim that you have met each learning outcome. Then make a list of the kinds or pieces of evidence that would back up such a claim. Then, again, find or develop for your teaching such evidence. You’ll find a form called ‘Evidence Planner’ in ‘Building your portfolio’ to aid that process, and Figure 4 shows that form in use.

You may want to devise your own form that will work better for you, or you may not like using forms, in which case you’ll need to find another way to work. If you come up with a good one, let me know for the next issue of the Portfolio:Guide.

However you do it, you need to have a method of collecting evidence; of knowing where it is; and of deciding which learning outcomes and which values it may demonstrate.
### Figure 3 Preliminary Evidence Organizer in use.

<table>
<thead>
<tr>
<th>Date of evidence</th>
<th>Course</th>
<th>Name/description of item</th>
<th>Source/author</th>
<th>Possible relevant outcomes (underline where filled)</th>
<th>Possible relevant values</th>
<th>Number of pages</th>
</tr>
</thead>
<tbody>
<tr>
<td>May 90</td>
<td>1st year 'Introduction to Psychology'</td>
<td>BA course document</td>
<td>Course leader</td>
<td>?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>May 90</td>
<td>Ditto</td>
<td>Course revision overheads</td>
<td>Self</td>
<td>1, 2, 4</td>
<td>1, 6</td>
<td>1</td>
</tr>
<tr>
<td>May 90</td>
<td>Ditto</td>
<td>Final student feedback (collated)</td>
<td>Students</td>
<td>3, 7</td>
<td>1, 2, 5, 6</td>
<td>3</td>
</tr>
<tr>
<td>May 90</td>
<td>Ditto</td>
<td>Marked essays</td>
<td>Students + self</td>
<td>3, 5, 6</td>
<td>all</td>
<td>Lots</td>
</tr>
<tr>
<td>June 90</td>
<td>Ditto</td>
<td>Final revision seminar plan</td>
<td>Self + students</td>
<td>1, 2, 3, 7</td>
<td>all</td>
<td>2</td>
</tr>
<tr>
<td>Feb 90</td>
<td>1st year 'Psychology research methods'</td>
<td>Student computer practical plan</td>
<td>Self + colleagues</td>
<td>3, 5, 6</td>
<td>1, 2, 4, 5</td>
<td>4</td>
</tr>
<tr>
<td>Mar 90</td>
<td>Ditto</td>
<td>1st year 'Introduction to Psychology'</td>
<td>Self</td>
<td>1, 2, 3, 7</td>
<td>1, 2, 3, 5, 6</td>
<td>2</td>
</tr>
<tr>
<td>June 90</td>
<td>Ditto</td>
<td>Results sheets for exam board</td>
<td>Self (using course form)</td>
<td>5, 6</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>July 90</td>
<td>N/A</td>
<td>Appraisal plan</td>
<td>Self</td>
<td>6, 7</td>
<td>All</td>
<td>3</td>
</tr>
</tbody>
</table>

### Figure 4 Evidence Planner in use.

<table>
<thead>
<tr>
<th>Item no.</th>
<th>Name/description of item</th>
<th>Source/author (if not self)</th>
<th>Possible relevant outcomes</th>
<th>Possible relevant values</th>
<th>Number of pages</th>
<th>Obtained?</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Course Guide</td>
<td>Course team</td>
<td>1</td>
<td>2</td>
<td>3, 4, 5, 6</td>
<td>pp. 3-6 (with course learning outcomes)</td>
</tr>
<tr>
<td>11</td>
<td>Previous lecturer's notes, to adapt or build on</td>
<td>KC</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>Lesson plans</td>
<td>self</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4, 5, 6</td>
</tr>
<tr>
<td>13</td>
<td>Seminar reading lists</td>
<td>self</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4, 5, 6</td>
</tr>
<tr>
<td>22</td>
<td>Lab worksheets</td>
<td>MR, modified by me</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4, 5, 6</td>
</tr>
<tr>
<td>23</td>
<td>OHPs for student presentations</td>
<td>Students</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4, 5, 6</td>
</tr>
</tbody>
</table>
Here is some further practical guidance on finding, making, collecting and sorting and collating evidence for your portfolio. You will probably find it useful to bookmark and return to this section as you build your portfolio:

- A blank piece of paper or screen on which you have to start a new piece of writing can be intimidating. An empty portfolio can have the same effect. So start collecting evidence!
- Keep your Portfolio:Guide with you each time you do any lesson planning, marking or giving feedback or administration.
- If in doubt, file possible evidence in your portfolio. It's easier and better to throw out evidence later which didn't quite do the job than try to find materials which you didn't file here at the time.
- Remember that each piece of evidence should help to show your competence as a teacher. All of the evidence you file will therefore be in some way personal to you – created by you, edited by you, or written by someone else for you or about your teaching. With a very few exceptions mentioned above, printed course materials and syllabuses, for example, rarely belong in your portfolio – unless of course you wrote them.
- Use the Evidence Organizer to log each piece of evidence in. This may feel a little over the top now, but it will help you keep track as your portfolio grows.
- Consider labelling each piece of evidence – for example, 'Outcome 3 item 7' – or perhaps just date them, or number them sequentially as you collect and file them.
- Don't worry if you can't decide exactly where every piece of evidence should go. Put it where it seems most logical to put it now – you can always move it later if necessary. If you really can't decide where it goes, put it into a section you call 'other useful evidence', or even in a separate box or file. But try not to let this section get too large! And try not to put things in this 'unsorted' file without at least making a note of the outcomes and values which they may demonstrate.
- If a piece of evidence seems to go equally well in more than one place, put a note about it in each of its other possible homes.
- Don't worry if the portfolio become rather untidy. It is a working document. You will move things around, add items, sometimes also reject items.
- As you build your portfolio, notice how much evidence is going into each section. Sections for some outcomes will naturally fill at different speeds – for example, you won't have anything on assessment until you've marked a set of student work. But some sections may stay very thin. If this happens, consider why:
  - Are you simply not doing any suitable work? In this case, negotiate with your course leader or mentor to make sure you have some teaching which can demonstrate this outcome;
  - Or perhaps you just don't see how your work is demonstrating that outcome. In this case, if it is possible, a conversation with another new teacher who is also building a portfolio may help. Swap portfolios; see what evidence other teachers are putting where; see if what they are doing has any implications for your portfolio.

But – before you go any further – if you haven't already done so, bring your Portfolio:Guide to life by starting to collect and sort and file evidence about your teaching work.

### 3.6 Showing principles and values in action

The idea of showing principles and values in action, underpinning practice, may be unfamiliar, although it is not always subtle or difficult. This section explores ways to do this.
Evidence about your teaching

Activity 1

(a) Which of the teachers below is, on the evidence presented to you, underpinning their practice with appropriate values as described in Section 2.2? In each case, what shows this?

(b) Note also what else they would have to do to persuade you that the value underpinned their practice.

I've shown my own response to the first example. You may find it useful to read and think about this before you work through the rest.

A An accountancy teacher writes a clear, accurate, comprehensive account of the equal opportunities policy of the university and of its implications for teaching.

My response

(a) The accountancy teacher clearly knows a lot about equal opportunities. However, we have no information here on how far this knowledge affects their practice.

(b) We would need to see them using this knowledge – for example, designing classes in a way which promoted equality of opportunity – and making it clear in their claim how equal opportunities considerations had informed their choice of methods.

Make a similar analysis for each of the following teachers.

B A biology teacher uses established texts on teaching and learning in higher education to help them to design their classes.

C A classics teacher plans, undertakes and reviews their teaching alone.

D A dentistry teacher has planned and prepared the first year of teaching in a new subject in great and careful detail, so that they can deliver it with little or no change for the next few years.

E An electronic engineering teacher is co-authoring a paper on a new approach to project work.

F A forestry teacher requires each student to relate the first assignment of the final year to the work placement from which they have just returned.

Comment

B

(a) The teaching work of the biology teacher is almost certainly underpinned by some understanding of how students learn, and perhaps by a commitment to scholarship.

(b) To show scholarship, we should additionally at least need to see the ideas from their reading applied in a critical way.

C

(a) The classics teacher does not seem to have embraced working with colleagues into their practice.

(b) To show collegiality in action, we should need to see evidence of the classics teacher working with other lecturers at some point in the planning, undertaking, assessing or reviewing of their teaching.

D

(a) The dentistry teacher, on the evidence, is either reluctant to reflect on their practice or reluctant to contemplate changing their practice on the basis of such reflection.

(b) Their current detailed teaching plans might be based on careful critical analysis of the teaching they have undertaken in previous years. If reflection is being claimed, then we would need to see evidence of this. And we would still be unhappy about an unwillingness to reflect on and change current practice – obtaining and using some student feedback during the current year would surely be good practice.

E

(a) The electronic engineering teacher, as well as showing collegiality, is very likely also to be
(a) The forestry teacher is showing a concern for students' development by valuing what each student brings to the course.

The kinds of judgement you were making here about values underpinning practice are the same kind your assessors will make (although they will have rather more evidence to go on).

You may find it useful to spend a few minutes thinking and making notes about how each of these examples relates to your own situation and practice. Don't worry about the subjects taught – concentrate on each teacher's behaviour.

Several points may have occurred to you as you worked through this exercise. The lessons for me (actually, the teaching points that I planned into this exercise) were:

1. Knowledge isn't enough, even when it has all the qualities demonstrated by the accountancy lecturer.

2. Even strongly espoused values (e.g., 'I really really really believe in equal opportunities') aren't enough. The values must demonstrably underpin your practice.

3. The values should wherever possible be clearly visible in your evidence alone – but you may need to annotate the evidence to make them clearly visible. For example, you could show how the principle or value led you to adopt the approach you adopted, and to reject others.

4. However, it is rarely enough to place the evidence before your assessor and ask them to make the judgement. You should make the case – that is, you should persuade, argue, show, indeed prove – that the evidence is there in your practice. Assessors need your help to make their judgements – they cannot know your work as well as you do.

In general, how can the values be demonstrated in what you do? Here are some general comments. Section 4 contains much more detailed guidance, outcome by outcome. The nature of the discipline which you teach will often have a major effect on the ways the values can be enacted. However, it will always be possible to show the values in action. Note that the examples below do not contain citations or other references to the literature. Such citations and references will be expected in your own claim (see the supplement Examples of Draft Claims and Tutorial Responses). Some sources are described in Section 6.1.

**Evidence for Value 1 – Understanding how students learn**

You could show how the theories and models of student learning you have met in the literature have led you to choose between alternative approaches to teaching or assessment.

'I was interested to read the accounts of deep and surface approaches to learning [citations omitted]. It helped me realize how some of my teaching was inadvertently pushing students towards surface approaches, and to see how to modify my teaching approaches to try to support and encourage deep approaches. I attach examples of my teaching approaches before and after the modifications.'

**Evidence for Value 2 – Concern for students' development**

You could show how your concern for the development of attitudes, skills and knowledge beyond the curriculum led you to particular teaching approaches, for example:

'I introduced short presentations by students, first to small groups and then to larger groups, to help them develop confidence and skill in speaking to an audience of peers. The support I offered included ...'
Evidence for Value 3 – Scholarship

Moving beyond Value 1, you could show your own pedagogic scholarship by developing and testing your own models and theories of teaching and learning as you proceed, in a way that is informed by the literature.

You should also show how you develop scholarship in your students, for example:

'I told students that, for the third assignment, they must find an author from the same period who was not on the reading list and see how well the critical frameworks I had been describing would apply.'

Evidence for Value 4 – Working with colleagues

Collegiality does not necessarily generate evidence in the way that other values do. You may need colleagues to confirm how you have worked with them.

For example, you might include a working group paper which you had co-authored or an exchange of memos or e-mails with colleagues which showed you working with colleagues to solve a problem or improve practice on the course.

Evidence for Value 5 – Equal opportunities

There may be two strands to this:

- showing how your teaching approaches and materials deliberately avoid exclusive or offensive language or stereotypes;
- showing how your teaching is actively inclusive and celebratory of difference.

For example, your seminar notes could show how you dealt with insensitive behaviour by a student, and the list of readings for a seminar paper could include work from a range of cultural perspectives.

Evidence for Value 6 – Reflection

You need to show that you reflect on your teaching. You could show how you continue to question your practice, explore the rationale for and the effects of what you do, bring in new ideas and approaches and evaluate their effects; for example:

'I was not happy with the small amount and low level of involvement in the seminar. I asked what the problem was and what we might do differently, and I managed to persuade them to give me a few suggestions. I'll keep the opening presentation much shorter next week, and get them working in pairs on a question – I'll see if that works any better.'

3.7 Judging evidence

You can only judge evidence thoroughly with the associated claim. Making a claim is explored in Section 4. Meanwhile here's a little guidance on judging evidence:

- In the early stages of portfolio-building, before you file each piece of evidence, ask yourself, 'Does this item really show that I have attained, or partly attained, or at any rate moved towards, this outcome and its underpinning values?' If it does, put it in. If it doesn't, don't. If it might but you're not sure, put it in anyway. You may be able to strengthen it later, or remove it when you find or produce stronger evidence.

- As soon as you have some evidence for one or more outcomes, start to write the claim (as discussed in Section 4). Writing the claim is the best way to see if you have sufficient and appropriate evidence. An early draft of a claim will thus steer your later collecting of evidence.

- In later stages, as your portfolio begins to fill up with evidence, ask yourself a slightly different version of this question about each new piece of evidence. Ask yourself, 'Does this item show that I have attained, or partly attained, or have moved towards, this outcome or value – and does it do so better than one of the
4 Making a claim for accreditation

To complete the assessment successfully, you need to show how you have attained each of the specified learning outcomes. You also need to show that you have done so in a way that is demonstrably informed by the underpinning values.

Table 1 summarizes which outcomes must be underpinned by which principles and values.

**Table 1 The relationship between outcomes and values.**

<table>
<thead>
<tr>
<th>Outcomes</th>
<th>1 Understand how students learn</th>
<th>2 Concern for students' development</th>
<th>3 Commitment to scholarship</th>
<th>4 Working with colleagues</th>
<th>5 Practise equal opportunities</th>
<th>6 Reflection</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Design teaching sessions</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 Use teaching and learning methods</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 Use assessment methods</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4 Monitor your work</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5 Keep records</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>6 Use coping strategies</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7 Reflect and plan CPD</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
</tbody>
</table>

This section takes you through each outcome in turn. For each outcome, it helps you to do two things:

A Decide which evidence you will use to show that you have attained the outcome

Section 3 offered general advice on collecting and generating evidence, and on sorting it under the appropriate outcome. The outcome-by-outcome advice offered in this section assumes that you have worked through Section 3, and have at least started to collect evidence and sort it by outcome.

The maximum number of pages (sides) of evidence for each outcome is specified.

B Write a claim that your evidence shows that you have attained this outcome, underpinned by the necessary principles and values

Making a claim involves spelling out how you have attained the outcomes and how your attainment of them has been underpinned by appropriate values.

Section 3.6 provided general guidance on showing principles and values in action in your work. A more detailed account for each outcome and its appropriate values follows.

The maximum length of each claim is specified.

Your assessors will assume that the claim under each outcome starts with a sentence such as this:

The evidence in the portfolio shows that I have attained Outcome (1), in that it shows how I have (designed teaching sessions from a course outline, document or syllabus, involving choosing teaching methods appropriate to the group of learners, the mode of study, the subject material, the resources available and the learning outcomes). It also shows how the specified values have underpinned my practice. In particular ... .

There is no need to write this every time.
4.1 Outcome 1 – Designing teaching sessions

Design teaching sessions from a course outline, document or syllabus. This involves choosing teaching methods appropriate to the group of learners, to the mode of study, to the subject material, to the resources available and to the learning outcomes.

A What kinds of evidence could you use?

Maximum 30 sides of evidence

Your evidence might include:

1. An annotated copy of the course outline, document or syllabus on which you based your detailed planning. (This will be one of the very few occasions on which you include in your portfolio material which you did not produce.) This document, together with your annotations, should show:
   - the course being taught (title and subject);
   - the overall course aim and learning outcomes;
   - the mode of study;
   - the kind of learners;
   - the year or level;
   - your role in and responsibilities for the course.

2. Some plans for teaching sessions. Each of these could usefully show:
   - the type of teaching session (e.g. lecture, seminar, laboratory);
   - learning outcomes to be achieved for that session;
   - the syllabus content to be dealt with;
   - an outline timetable for the teaching session.

B How will you claim that your evidence shows that you have, in a way that is underpinned by the necessary principles and values, attained this outcome?

Maximum claim length: 750 words

Value 1 – How students learn

Many of the models and theories described in the Reader for H851, and elsewhere in the literature on teaching and learning in higher education, have implications for the design of teaching sessions and for the choice of teaching methods. You will also build your own models and theories about learning and teaching. (Indeed, you will already have developed such models and theories, although you may not yet have made them explicit or tested them systematically.)

You could show your understanding of how students learn by justifying your design of teaching sessions and choice of teaching methods with explicit reference to theories or models of student learning.

For example, your class plan shows:
- you explaining an approach or technique to the students;
- them applying this to an example or question;
- you asking some of them to tell you and each other what conclusions they
them reviewing their work in the light of your explanation, exploring any differences between your approach and theirs, and then starting to plan how they will tackle the private study task.

Your claim on this shows how the plan was intended to take students through a cycle of learning in the class, with reference to Kolb's or a similar learning cycle model.

**Value 2 – Concern for students' development**

In your design of teaching sessions, you might show your concern for student development, in part, by focusing on and making use of the needs and goals that students bring to the class.

For example, in the first meeting of the class you ask students to spend a few minutes noting down the three questions they hope that, above all, the course will answer. You collect these in, and use them as far as you can to adjust the emphasis and organization of the classes.

**Value 5 – Practising equal opportunities**

You can show equal opportunities practice in your design of teaching sessions both by what you choose to do and by what you avoid doing. You can show how your examples and your methods are inclusive, and you can show how you avoid discriminatory or excluding practice and language.

For example, early in the course you have noticed that some members of the group are reluctant to participate. Your plan for later seminars shows sensitivity to this by including pair work in which people can choose the partner with whom they work.

Also, the examples you use are not unnecessarily from a single gender, ethnicity or sexual orientation, avoid stereotyping, and acknowledge diverse abilities.

So, a claim for Outcome 1 might start like this:

Evidence 1.1 shows the original syllabus and the overall semester programme I planned from it; 1:2 shows detailed plans for four classes; 1:3 shows student feedback which indicates that they found the teaching methods appropriate (though at times a little more challenging than they expected).

Why do I claim that the teaching methods were appropriate to the learning outcomes? The learning outcomes required students to be able to apply price theory to particular industrial and commercial situations. I broke this down to show that there were really four main steps in applying price theory. I thought of teaching these steps one at a time and then putting them together at the end. However, I was impressed by the view that many students need to see the big picture before they can start to learn the details [citations omitted]. (I certainly remember that problem as a student, not seeing the wood for the trees in a lecture.)

So I approached it top down. We looked at real situations, and I asked them to analyse what was going on. Then I introduced price theory as a way to solve a real problem, rather than as a pile of wooden blocks and no picture of what to build. The students said they liked it, although it made them think and work in class more than they were used to. Also, one example was far too difficult and I had to simplify it on the spot before they became too upset. Most of them did well on the first piece of assessed coursework, although a few are clearly struggling ...

When you have completed this section of your portfolio, make these checks:

1. Each of items 1c–1j in the Portfolio Assessment Form is explicitly addressed.
2. Every item of evidence is referred to in the claim.
3. There are no more than 30 pages of evidence.
4.2 Outcome 2 – Using teaching and learning methods

Use two appropriate teaching and learning methods from: making presentations (e.g. lectures, demonstrations); facilitating group learning (e.g. through seminars, discussion groups, projects); working with individual learners; facilitating practicals or laboratory classes, also including evidence of using appropriate learning technologies.

A What kinds of evidence could you use?

Maximum 30 sides of evidence

Your evidence must show that your teaching has been observed by a peer and that you have reflected on that observation. Additionally, your evidence might include:

Teaching method 1
- Annotated lesson plans to show that you structured your teaching to introduce a variety of appropriate methods and material.
- Examples of tasks and activities that you required students to engage in.
- A short description of the learning technologies that you used, showing how they supported which learning outcomes. These might include:
  - computer-based learning;
  - web-based learning;
  - videos;
  - open learning texts;
  - workbooks;
  - activity sheets;
  - overhead projection transparencies;
  - structured library assignments.

Teaching method 2
Find similarly varied evidence to show how you have achieved Outcome 2 in a different type of teaching. So, if you chose lecturing as your first example, you may this time choose running seminars or projects.

B How will you claim that your evidence shows that you have, in a way that is underpinned by the necessary principles and values, attained this outcome?

Maximum claim length: 750 words

Value 1 – How students learn
You can readily show your concern for how students learn in action in your teaching, as long as you can produce for your portfolio a detailed account of what you do in a class.

You will already have shown how this value informs the planning of a class in Outcome 1. Here you have to show this value in action in the details of the class. For example, how does your class show your students having opportunities to be active, to test their ideas, to see their progress?
which you respond to questions in your classes. You also need to justify, with
reference to models or theories, your different kinds of responses to student questions.
Consider your responses in terms of their contribution to the learning of an individual
student and the whole class at a particular moment.

Value 2 – Concern for students’ development

Your conduct of a class should, this value suggests, go beyond the immediate learning
outcomes of the class and also support students’ broader personal and professional
development, for example their skills of communication and co-operation.

For example, the practical class involves a fairly standard experiment. You ask them
to write it up as a critical evaluation of the experiment, rather than the usual
formulaic report. You thus encourage them to develop an additional valuable skill.

Value 5 – Practising equal opportunities

The relevance of equal opportunities to class planning was discussed earlier. You are
trying to develop an open, knowledgeable and respectful approach on the part of
yourself and your students which will allow everyone to participate.

For example, in a seminar you notice that two students dominate the discussion stage
and alienate some students by their views on gender and class. You discuss with the
whole group the importance of adhering to agreed ground rules on equal
opportunities and ask others to say how well this is working. You create space for
others to open up and say how the prejudicial behaviour affects them. You elicit
responses from the two students, and ask the group to re-negotiate and re-affirm the
ground rule.

Value 6 – Reflection on professional practice

Reflection is essential to the improvement of teaching. You show your continual
reflection on your practice through examples of reviews of your practice and what you
did as a result.

For example, you question your teaching methods as the course progresses.
Subsequent lesson plans show your teaching changing in the light of the answers to
these questions. Your claim makes your learning from your evidence-based reflection
explicit. You might make explicit use of a structured process of reflection based on a
learning cycle (see Kolb et al., 1991, in Section 6.1).

So a claim for Outcome 2 might start like this:

In reflecting on my first example of a method that I use a lot in my teaching –
lecturing – I realize that I do think of a lesson plan in terms of what students will
learn as well as what I do to help them.

Initially I tended to be focused entirely on delivering content. I began to discover
through seminar discussions and written assignments that students were not
always picking up the structure and principles of the course.

My plans show how I break the lecture to set a question for small groups, pairs
and individuals to engage in. I have had to work at pitching the tasks at the right
level and scale, so that I can manage quite large group activities and still get
back to the lecture.

The set of handouts, which are an example of my use of learning technology, are
based on an idea of gapped handouts I got from [citation omitted]. They helped
me manage these sessions during lectures.
When you have completed this section of your portfolio, make these checks:

1. Each of items 2c–2j in the Portfolio Assessment Form is explicitly addressed.
2. Every item of evidence is referred to in the claim.
3. There are no more than 30 pages of evidence.
4. The claim is no longer than 750 words.
5. Two teaching sessions were observed and reported.
6. Each session involved appropriate use of learning technologies.

4.3 Outcome 3 – Using assessment methods

Mark or grade, and give feedback on, student work.

A What kinds of evidence could you use?

Maximum 30 sides of evidence

Your evidence might include a short description of at least one assessment method you have used that is appropriate to the learning outcomes that your students are aiming to achieve. This may be, for example, your assessment of a written essay assignment, a critique of a project or an assessment of a seminar presentation.

You should also provide evidence to show how you did the following tasks:

1. Grading and marking – for example, a copy of a marking sheet showing how marks and grades are to be distributed to reflect agreed assessment criteria and an analysis of the resulting set of marks that you awarded showing its distribution across grades.

2. Giving feedback to students – for example, some written comments that you have attached to returned work.

B How will you claim that your evidence shows that you have, in a way that is underpinned by the necessary principles and values, attained this outcome?

Maximum claim length: 750 words

Value 1 – How students learn

Your feedback to students should be designed, in an informed and reasoned way, to support their learning. Your evidence must show how you have done this, and on which models of student learning you have based your approach to assessment.

For example, you have found, by asking your students about assessment, that their main concern is how to get the mark or grade that they want. Accordingly, alongside an indicative mark or grade in each piece of work, you tell them what they would need to do next time to gain a higher mark or grade for such a piece of work.

Value 2 – Concern for students’ development

Assessment may not always feel very developmental to students, particularly if you...
assessment or given feedback to address their broader personal and professional development as well as the outcomes of the particular course or unit.

For example, the assessment of the course is, in part, by a small project or dissertation. You encourage students to define their own projects, to reflect their own needs and interests, within a clear framework of formal requirements.

Value 5 – Practising equal opportunities

Assessment is intended to discriminate between those who have attained the outcomes of the course to varying extents. It may also discriminate unintentionally on other criteria, such as gender. Show what steps you have taken to prevent or reduce this.

For example, you have become aware in marking the work of a couple of students that they may be dyslexic. You seek professional guidance from your student services unit, which provides the students with an opportunity for assessment and special tuition.

Before first-year exams you clarify with your course team leader what allowances can be made for dyslexic students undergoing predominantly written assessment.

Value 6 – Reflection on professional practice

The assessment of your students gives you feedback on your teaching. Show how you have heard and learned from this feedback, and adapted your practice accordingly.

For example, you analyse students’ answers to questions, identify the main types of errors and confusions, work with your students to identify how these confusions and errors arise, and modify your teaching accordingly.

So a claim for Outcome 3 might start like this:

I have included an assignment sheet that I have developed to solve the problem that students had when I first set the task of writing up practicals. Some wrote far too much detail whilst showing little understanding of the scientific purpose of the experiments they described. Others failed to record clearly what they had observed and measured.

My assignment sheet spells out the learning objectives of this type of written task, and then shows in detail the criteria that students have to meet to achieve a satisfactory grade. It also points out what more could be done to achieve a higher grade.

When you have completed this section of your portfolio, make these checks:

1. Each of items 3c–3h in the Portfolio Assessment Form is explicitly addressed.
2. Every item of evidence is referred to in the claim.
3. There are no more than 30 pages of evidence.
4. The claim is no longer than 750 words.

4.4 Outcome 4 – Monitoring your work

Monitor and evaluate your own teaching, using self, peer and student feedback.

A What kinds of evidence could you use?
Your evidence might include:

1. your own teaching diary or log recording issues, problems and solutions that you have reviewed and acted on;
2. notes made by you before, during and after a particular teaching session;
3. copies of student feedback from particular sessions or a whole course;
4. written feedback from a colleague who has observed you at work;
5. notes that you prepare for an appraisal interview with your manager or head of department, and any summary of feedback that they provide for you. (Even if such a scheme does not exist in your organization, it would be useful to request an informal appraisal interview, to review your progress.)

B How will you claim that your evidence shows that you have, in a way that is underpinned by the necessary principles and values, attained this outcome?

Maximum claim length: 750 words

Value 1 – How students learn

Monitoring and evaluating your teaching provides an opportunity for you to develop further your understanding of how your students, this year, in this class, learn. The guiding principle here is to ask questions and collect information that will enable you to understand what is going on and, where necessary, change what you do.

Evaluating your teaching; increasing your understanding of how your students learn; applying that understanding to changing your practice; evaluating the effects of the changes you make – this may all sound like a kind of applied research. You may be able to apply some of your research skills to your teaching.

The need to monitor in order to understand has implications for the way that you monitor.

Consider this question:

‘Do you agree – feel neutral – disagree that the pace of the class was about right?’

You may find that a third of your students agreed, another third was neutral and another third disagreed. This doesn’t tell you whether you should speed up or slow down to meet the needs of those who disagreed. Nor does it tell you whether you should change the pace of the class at all!

Show this value in action for this outcome by asking questions the answers to which can enhance your understanding and guide you on action.

Value 2 – Concern for students’ development

Monitor all your teaching work with regard to its effect on student development. Are your students getting an optimal mixture of comfort and stress, of support and challenge? Do they feel valued, and can they see themselves developing?

For example, you set aside some time during group tutorials halfway through a module or unit to encourage students to give you feedback on some key topics concerned with their development. These might include:

What different perspectives have they developed about the subject?
What new skills have they developed?
What learning activities have they found most useful in developing these perspectives and skills?
What was least useful?
**Value 3 – Scholarship**

Be scholarly in your monitoring and evaluation. Use evaluation methods which have some backing in the literature, rather than simply making them up, especially if your disciplinary background has not given you training in such skills as survey design. Analyse your results systematically so as to extract valid information. Alternatively, if there simply isn’t time to develop and apply the necessary skills, be clear about the limits of your monitoring, and don’t claim too much for your results or use them overconfidently or inappropriately.

For example, you might explore some of the literature on the collection and use of student feedback on teaching, for example Beaty (1997) and Diamond (1998) in Section 6.1. As a consequence, you may redesign some questions on the standard form you have so far used, or experiment with new ways which do not involve forms.

**Value 4 – Working with and learning from colleagues**

Seek out monitoring and evaluation methods used by colleagues. If the local culture allows this, share and review your monitoring and evaluation results with colleagues. They may reciprocate, and useful learning may follow for you and your colleagues.

For example, you may want to show what use you have made of co-teaching or observation and feedback on your teaching.

**Values 5 and 6 – Practising equal opportunities and reflection on professional practice**

Monitor that your work as a teacher conforms to the equal opportunities policies of your institution.

Equality of opportunity cannot be fully ensured by adherence to policies or checklists. However:

- test your planning, your teaching and your assessment against equal opportunities policies and checklists to ensure that you avoid bad practice;
- beyond that, explore what it means to provide equality of opportunity for all students for whom you have some responsibility. Chronicle the steps you take to implement the conclusions you reach;
- monitor the effects of what you do, reflect on the reasons why it works or doesn’t and change your plans and your practices accordingly. And continue to do so!

So a claim for Outcome 4 might start like this:

I have attached the notes that my colleague gave following her observation of a lecture and associated seminar that I taught. She did this as a part of our Certificate programme and we had already agreed on general criteria that would be used on the course. In addition I talked through one or two particular points that I wanted feedback on. These included:

In the lecture:
- Do I give students a sufficiently clear overview of the learning objectives and the content of the session when I start?
- Do I signpost clearly enough when I am moving from one step in the argument to the next?
- Do I summarize key learning points at the end?
- Do I create a rapport with the students and read their level of interest and engagement in what I am saying?

In the seminar:
- Do I leave responsibility for the delivery of the paper with the two students acting as seminar leaders?
- Do I use my interventions to facilitate maximum involvement by all participants?

The written feedback from my colleague was a source of great encouragement to me but I also identified that I still need to make this—change.
When you have completed this section of your portfolio, make these checks:
1. Each of items 4c–4k in the Portfolio Assessment Form is explicitly addressed.
2. Every item of evidence is referred to in the claim.
3. There are no more than 30 pages of evidence.
4. The claim is no longer than 750 words.

4.5 Outcome 5 – Keeping appropriate records

Keep appropriate records of your teaching support and academic administrative work.

A What kinds of evidence could you use?

Maximum 20 sides of evidence

As part of an academic community, whether at course team level or the whole organization, you have a range of tasks to perform which support teaching and maintain administrative efficiency and effectiveness.

The evidence for this outcome will vary widely depending on the way your organization devolves such tasks. Your evidence might include:
1. The appraisal interview notes submitted in support of Outcome 4 which may include an assessment of your contribution to committees and your handling of administrative tasks.
2. A paper you have written for a course team.
3. A copy of student attendance records and marks.
4. Plans that you have drawn up to achieve a particular task at work, such as preparing new course material or planning a student visit.

B How will you claim that your evidence shows that you have, in a way that is underpinned by the necessary principles and values, attained this outcome?

Maximum claim length: 500 words

Support and administration may sound to be value-free, or at any rate value-light, activities. In fact, you should show how reflection informs these activities.

Value 6 – Reflection on professional practice

Much teaching support and administration involves little more than identifying local rules and procedures and following them. However, there is usually some freedom over how to interpret procedures. When this is so, reflect a little on how the procedures may best be followed. You could explore what the procedures are designed to achieve, and act in support of these goals; you could explore how you could follow the procedures to meet your goals of efficient use of your time and avoidance of unnecessary conflict or difficulty.
You will also need to develop some local procedures to support and administer your own teaching. Again, a little reflection on what your procedures are designed to achieve can lead you to better and more appropriate procedures.

For example, principles for the design of your local systems might include:
- avoid, wherever possible, manual copying of information;
- ensure that you have a back-up in a different place;
- start by asking, 'Who will want this information, when, and in what form?';
- make the system flexible – administration and information requirements change.

So a claim for Outcome 5 might start like this:

The evidence of my involvement in the course team can be found in a report I put together on ways of improving work placements in the third year. There have been some lapses in managing placements in previous years, so I was asked to review the arrangements in the light of our redefined policy on using work placements.

In particular we have now got a database of employers and better record-keeping on how students fare with each of them. I also planned the revision of the Work Placement Guide and collaborated with two colleagues to get it re-written, designed and printed. The minutes of our planning meetings and sample pages of the new Guide are attached as evidence.

When you have completed this section of your portfolio, make these checks:

1. Each of items 5c–5d in the Portfolio Assessment Form is explicitly addressed.
2. Every item of evidence is referred to in the claim.
3. There are no more than 20 pages of evidence.
4. The claim is no longer than 500 words.

4.6 Outcome 6 – Using coping strategies

Develop personal and professional coping strategies appropriate to the constraints and opportunities of your institutional setting, to manage adequately your time and operate successfully within available resources.

A What kinds of evidence could you use?

Maximum 20 sides of evidence

We all have our own way of coping. The evidence for this outcome will be very much self-reported and individual to you. Brief written evidence might include:

1. an analysis of a working week or fortnight to show how time was spent and an assessment of how it might be better spent to meet your professional objectives;
2. a written summary of a work overload problem and the way you negotiated to resolve it;
3. an account of a way in which you have made the most of scarce teaching resources;
4. a record of any steps you have taken to monitor and reduce stress – for example...
B How will you claim that your evidence shows that you have, in a way that is underpinned by the necessary principles and values, attained this outcome?

Maximum claim length: 500 words

Coping is not a particularly academic, scholarly or indeed value-laden business. However, two values should be shown here – collegiality and reflection.

**Value 4 – Working with and learning from colleagues**

Your colleagues are part of what you have to cope with. They should also be part of your coping strategy. Show how you support them, seek their advice, work with them, gain support from them.

For example,

- How have you identified the sources of stress that impinge on you? What were they?
- Which of your colleagues may be able to help you to understand or reduce each major source of work stress? How have you worked with them to achieve this?
- Which if any of your coping strategies have involved your colleagues?

**Value 6 – Reflection on professional practice**

You need to show how you have reflected on the sources of work-related stress, and on the effectiveness or otherwise of the coping methods you have adopted.

For example, you may reflect on the extent to which your current coping strategies are working. What image do you draw on to picture yourself as a member of the teaching profession? If you have worked in other jobs, compare your experience of coping with pressures in each.

Ask yourself where you aim to be in five years' time; is your current way of coping with and responding to demands going to ensure that you get there or do you need to address particular problems?

**So a claim for Outcome 6 might start like this:**

The evidence I have submitted illustrates my particular approach to coping with the pressures of the job. I am not easily stressed by things that are within my own control; for example, I found planning and writing my first term’s teaching material hard work but very satisfying. I was lucky because I had a lot of material from my PhD, which I could draw on directly.

But I have found that the unexpected demand to fit in another seminar group and develop a course in an area that is fairly new to me has caused great pressure. The documents that I attach as evidence show the steps I took to discuss this with the Course Tutor and a colleague who has been particularly helpful. I have also included the planner that I used as a discipline and an incentive to get the extra work ready on time.

When you have completed this section of your portfolio, make these checks:

1. Each of items 6c–6f in the Portfolio Assessment Form is explicitly addressed.
2. Every item of evidence is referred to in the claim.
3. There are no more than 20 pages of evidence.
4. The claim is no longer than 500 words.
4.7 Outcome 7 – Reflecting and planning CPD

Reflect on your own personal and professional practice and development, assess your future development needs, and make a plan for your continuing professional development.

A What kinds of evidence could you use?

Your evidence might include:

1. a summary of what you have learned from your teaching work and your capabilities and deficiencies as a teacher;
2. a particular episode in your teaching record which shows a turning point in your personal progress;
3. a plan for attending staff development events on teaching within or outside your organization;
4. a reading programme to increase your understanding of teaching and learning in higher education;
5. a brief commentary on several papers which have influenced your thinking on your continuing professional development.

B How will you claim that your evidence shows that you have, in a way that is underpinned by the necessary principles and values, attained this outcome?

This may feel a little complicated. We've been asking you throughout to show how reflection underpins pretty much everything you do as a teacher, and now we're asking you to describe how each of the values underpins your ability to reflect.

I believe that reflection is the most effective single method for improving as a teacher (or for that matter as a member of any other profession).

And the values can – must – inform the way you reflect.

Value 1 – How students learn
You should reflect in order to test and extend your understanding of how students learn. I return to this under Value 3, scholarship.

For example, your reflection should be informed by three main sources:

- by questions, about your teaching and about your students' learning; questions which you really want to answer, because the answers will help you to confirm or improve your practice;
- by the ideas, theories, models and methods in the literature on teaching and learning in higher education, gained from the H851 Reader and elsewhere;
- by data gained from your own work as teacher.

Value 2 – Concern for students' development
Similarly, you should reflect in order to test and extend your understanding of student
For example, reflect on the learning relationships that you are developing with students. Can you see signs that some or all are beginning to share in your enthusiasm for your subject? What is it that you do which fires their imagination about being a member, a practitioner of your discipline?

**Value 3 – Scholarship**

As discussed earlier, your reflection should be scholarly; with regard to the teaching and learning methods you choose, the ways in which you use or adapt these methods, the conclusions you draw and the use you make of these conclusions in your work. Scholarly reflection of the type suggested here can form a basis of a further form of scholarship; your own conference presentations and published papers on pedagogy, generic or in the pedagogy of your discipline.

For example, assess to what extent you draw on published scholarship on teaching and learning, generic or in your discipline. What type of material has been of most help? Which writers have contributed most to your development? Why?

**Value 4 – Working with and learning from colleagues**

Reflection need not be a solitary activity. You should use appropriate colleagues at each stage of the process of critical, informed, understanding- and action-oriented reflection on your teaching. This co-operation can be through means as simple as a conversation, or by exchange of an commentary on academic papers, by local meetings or seminars, by collaborative investigation and publication, by running a conference – one or more of many forms of collegiality might be appropriate in support of your reflection.

For example, you might reflect on the extent to which you have been helped and supported by different colleagues as you gain experience as a teacher. What is it in the relationship that has encouraged and facilitated your learning?

Conversely, you may have some colleagues with whom you find it hard to work. Reflect on what it would take to confront this problem in a constructive way.

**Value 5 – Practising equal opportunities**

Reflection should embrace the extent to which equal opportunities considerations have informed your teaching. Insofar as reflection is a collaborative rather than a private activity, the ways in which you collaborate in reflecting should be supportive. Your partners in reflection can usefully be chosen in part for their diversity, helping you to see new issues and to take, perhaps, unfamiliar approaches to reflection and analysis.

For example, I suspect that we all have elements of prejudice in us which can be expressed intentionally or unintentionally in discriminatory words or actions. (The lack of intention is, of course, no defence; the impact of discrimination will still be felt by students or colleagues.)

You may need to ask yourself in what ways your own prejudice has become more clear to you as your teaching has progressed. In what ways have you behaved differently to minimize the risk that prejudice is expressed through discriminatory behaviour?

**Value 6 – Reflection on professional practice**

Reflection on reflection? You may be feeling, however you choose to express it, that I have gone too far! If you do, I disagree with you. It is possible, and can be very valuable, to reflect on reflection, as on any other professional activity. Your reflection on reflection can increase your understanding of methods of reflection and guide you to find more effective ways to teach, for example:

- What questions have you asked yourself in your process of reflection?
- Were the answers useful, illuminating, productive of understanding and guidelines for action?
- What kinds of questions were particularly useful?
What methods of analysis were particularly useful, and, again, what further such methods could you seek for development?

So a claim for Outcome 7 might start like this:

My main piece of evidence to support this outcome is the teaching diary that I have kept throughout the time I have been preparing this portfolio. At first I doubted the value of filling in entries and was not altogether sure what was worth including.

However I compared notes with X, who is also doing the course, and we discussed the diary idea with our head of staff development. This proved very helpful and I learned to use the diary for a range of reflective ideas such as notes from my course work reading, accounts of problems in my teaching and working notes on how to solve them.

I have selected the items of evidence from the diary, to illustrate how valuable the process has been for me. I plan to continue the diary after I have finished the course. I am also using the same technique to shape the next stage in my research ...

When you have completed this section of your portfolio, make these checks:

1. Each of items 7c–7k in the Portfolio Assessment Form is explicitly addressed.
2. Every item of evidence is referred to in the claim.
3. There are no more than 20 pages of evidence.
4. The claim is no longer than 500 words.

4.9 The overall assessment of your portfolio

You are approaching the end of the course. You have:
- collected evidence;
- filed each piece under the appropriate learning outcome;
- checked the evidence for appropriateness;
- sifted the evidence to ensure that it is as strong as it can be;
- identified any gaps in your evidence and filled them; and
- written a claim that shows how you have attained the learning outcomes in a way that is underpinned by the values.

We suggest one more step before you send your portfolio to your assessors. Assess the portfolio yourself. Put yourself in the place of an assessor. Assess your claim, identify any remaining gaps in the evidence, and take the necessary steps to fill them.

How?

Work through the guidance notes that have been prepared for your assessors (they are in the Assessor Guide booklet). It is important that you understand how the assessors will reach their decision.

4.10 What to do next?

Depending on the results of your assessment of how well you have done, you should follow one of three routes:

A. Your review may have persuaded you that you had demonstrably attained the outcomes, in a way that was demonstrably underpinned by the necessary values. In this case you should continue to submit your portfolio by the next assessment.

B. You may feel that you have achieved the outcomes, in a way that was informed by the underpinning values. However, you may feel that the evidence for one or
generate the evidence. Depending how long you think this will take you, you may
decide to submit your portfolio at a subsequent assessment point.

C You may feel that you have not yet fully attained one or more of the outcomes, or
not yet done so in a way that was fully informed by the necessary underpinning
values. In this case you will need to undertake further teaching, or perhaps to
provide simulated evidence for part of an outcome. Depending on the scale of this
necessary additional work, you may well decide to choose a later date to submit
your portfolio for assessment.

5 Submitting your portfolio for assessment, and what
happens next

5.1 Administrative arrangements

You should make two copies of your completed portfolio. Assemble these copies using
the dividers provided, and put them into the return portfolio binders provided with
the course materials. Post these using the return mailer provided.

5.2 Further studies in higher education practice

When you have successfully completely completed H851 or HH851, Teaching in Higher
Education, and thus achieved Associate Teacher status, I hope that you will want to
continue to develop as a teacher.

The next Open University course is H852/HH852, Course Design in Higher Education.
As well as course design, H852 deals with the development and use of a wider range
of teaching and assessment methods.

H852/HH852 is currently being designed to lead to SEDA Accredited Teacher status.
When the accreditation requirements of the Institute for Learning and Teaching are
clarified, H852/HH852, like H851, will be modified as necessary to reflect these.

You will be sent details of H852/HH852 when you complete H851.
Successful completion of these two courses leads to the award of Postgraduate
Certificate in Teaching and Learning in Higher Education.

6 Further information

6.1 Publications and other sources

Seldin, P. (1997) The Teaching Portfolio: a practical guide to improved performance and
promotion/tenure decisions, Bolton (MA), Anker, ISBN 1 882982 15 0.
This American text gives practical advice and guidance on the construction and use
of portfolios of evidence of teaching, with examples.

Ballantyne, R., Bain, J. and Packer, J. (1997) Reflecting on University Teaching:
47287 1.
A rich source of lecturers’ thoughtful accounts of teaching, across a wide range of
disciplines and teaching topics.

Teaching in Higher Education: theory and evidence, Milton Keynes, The Open
This chapter of the Reader for H851, *Teaching in Higher Education*, discusses how teachers change and develop in their professional practice and explores the role of reflection therein.


Practical guidance on reviewing your teaching and learning, and thus on developing information and ideas which can form part of your portfolio.


This account of a cycle of learning through the analysis of experience underpins this course.


A valuable short guide to reflecting on, monitoring and improving teaching.


This large and thorough book contains many ideas and resources for evaluating courses.

*The New Academic* (published by the Staff and Educational Development Association).

This termly magazine includes practical articles on teaching, learning and assessment.

*Innovations in Education and Training International* (published by the Staff and Educational Development Association).

This refereed quarterly journal publishes research-based papers on many aspects of higher education practice.

*Journals on teaching and learning in your discipline* A host of these can be found through *Deliberations*, in itself a valuable source of materials on learning, teaching and assessment in higher education, at http://www.lgu.ac.uk/deliberations/

**Conferences and workshops on higher education practice**

An up-to-date list of these is maintained by the Staff and Educational Development Association on their website at http://www.seda.demon.co.uk.

**6.2 The accreditation of teachers in higher education**

The idea that teachers in higher education should be accredited as teachers is relatively recent. Academics are usually highly qualified in the discipline or profession which they will teach; much more rarely in the business of teaching. This contrasts strongly with school teaching, where a professional qualification in teaching is a requirement, and with teaching in further education, where many teachers hold a further education teaching certificate.

Initial training in teaching has been provided by some universities and colleges for many years. Courses range from one-day introductions to teaching to substantial part-time programmes. However, since 1992, a steadily growing number of universities and colleges have been developing and running courses to common national standards, developed by the Staff and Educational Development Association (SEDA). H851, and its successor H852, have been designed to these standards.

You will already have seen that the SEDA standards specify the overall learning outcomes of a course to train higher education teachers, along with the principles and values which must demonstrably underpin the practice of the accredited teacher (Section 2). SEDA undertakes a formal process to 'recognize' courses. Courses are
of their institution, and have an appeals process.) When I was writing this (June 1998), 74 courses were recognized by SEDA or registered and preparing for SEDA recognition, most in the UK but also in Australia, New Zealand and Singapore, and over 700 teachers had been accredited. Further information is available at the SEDA website, http://www.seda.demon.co.uk.

The UK National Committee of Enquiry into Higher Education (the Dearing Committee) in 1997 recommended that all new full-time teachers in higher education should be trained through accredited programmes. A new Institute for Learning and Teaching (ILT) was proposed to undertake this work. A committee chaired by the former Vice-Chancellor of Oxford Brookes University, Professor Clive Booth, undertook the development of a new national higher education teacher accreditation framework for the ILT. The draft framework is available via the website of the Committee of Vice-Chancellors and Principals (CVCP) at http://www.cvcp.ac.uk/boothfin.html.

The current (June 1998) ILT accreditation proposals say that those who successfully complete programmes similar in outcomes to the proposed ILT standard should be awarded, for an interim period, the appropriate grade of ILT membership. The outcomes of H851 are indeed similar to the proposed ILT standard, which means that the professional qualification you gain from this course should gain you Associate Membership of the ILT.
Building your portfolio
Preliminary Evidence Organizer

Please make further copies for your use. A partially completed example is shown in Section 3.5.

<table>
<thead>
<tr>
<th>Date of evidence</th>
<th>Course</th>
<th>Name/description of item</th>
<th>Source/author</th>
<th>Possible relevant outcomes (underline where filed)</th>
<th>Possible relevant values</th>
<th>Number of pages</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Evidence Planner

Please make further copies for your use. A partially completed example is shown in Section 3.5.

<table>
<thead>
<tr>
<th>Item no.</th>
<th>Name/description of item</th>
<th>Source/author (if not self)</th>
<th>Possible relevant outcomes</th>
<th>Possible relevant values</th>
<th>Number of pages</th>
<th>Obtained?</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Outcome 1 Evidence Organizer

Use this form to log the evidence you may include for Outcome 1 and for the related professional values. (Values which do not apply to a particular outcome are greyed out.)

Make and use further copies of this form if you find it helpful.

Outcome 1 — Design teaching sessions from a course outline, document or syllabus. This involves choosing teaching methods appropriate to the group of learners, the mode of study, the subject material, the resources available and the learning outcomes.

<table>
<thead>
<tr>
<th>Name or description of item</th>
<th>Date</th>
<th>Source/author</th>
<th>Relevant professional values</th>
<th>No. of pages (max. =30 overall)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>1 How students learn</td>
<td>4 Working with colleagues</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2 Concern for students' development</td>
<td>5 Equal opportunities</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>3 Commitment to scholarship</td>
<td>6 Reflection</td>
</tr>
</tbody>
</table>

Total = not more than 30
Outcome 2 Evidence Organizer

Use this form to log the evidence you may include for Outcome 2 and for the related professional values. (Values which do not apply are greyed out.)

Make and use further copies of this form if you find it helpful.

Outcome 2 – Use two appropriate teaching and learning methods from: making presentations (e.g. lectures, demonstrations); facilitating group learning (e.g. through seminars, discussion groups, projects); working with individual learners; facilitating practicals or laboratory classes. You must also include evidence of using appropriate learning technologies.

<table>
<thead>
<tr>
<th>Name or description of item</th>
<th>Date</th>
<th>Source/author</th>
<th>Relevant professional values</th>
<th>No. of pages (max. = 30 overall)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>1 How students learn</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2 Concern for students' development</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>3 Commitment to scholarship</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>4 Working with colleagues</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>5 Equal opportunities</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>6 Reflection</td>
<td></td>
</tr>
</tbody>
</table>

Total = not more than 30
Outcome 2 Observation Schedule

You need to show that you have had two teaching sessions observed.

Please use this form to show how you have planned this observation; how the observation took place; what the observer reported; what you planned to do following the observation, and what you did in response to the observation.

It's fine to attach separate reports in each case as part of your 30 pages of evidence.

Name of observer:

1 Observation planning meeting Date:

Agreed observation schedule or main issues (attached if a separate document)

2 Observation Date:

Observation report (attached if a separate document)

3 Post-observation meeting with observer Date:
   Notes of meeting and actions agreed (attached if separate document)

4 Actions taken (attached if separate document)
## Outcome 3 Evidence Organizer

Use this form to log the evidence you may include for Outcome 3 and for the related professional values. (Values which do not apply are greyed out.)

Make and use further copies of this form if you find it helpful.

Outcome 3 – Mark or grade, and give feedback on, student work.

<table>
<thead>
<tr>
<th>Name or description of item</th>
<th>Date</th>
<th>Source/author</th>
<th>Relevant professional values</th>
<th>No. of pages (max. = 30 overall)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>1. How students learn</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2. Concern for students' development</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>3. Commitment to scholarship</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>4. Working with colleagues</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>5. Equal opportunities</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>6. Reflection</td>
<td></td>
</tr>
</tbody>
</table>

Total = not more than 30 pages
Outcome 4 Evidence Organizer

Use this form to log the evidence you may include for Outcome 4 and for the related values.

Make and use further copies of this form if you find it helpful.

Outcome 4 – Monitor and evaluate own teaching, using self, peer and student feedback

<table>
<thead>
<tr>
<th>Name or description of item</th>
<th>Date</th>
<th>Source/author</th>
<th>Relevant professional values</th>
<th>No. of pages (max.=30 overall)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 How students learn</td>
<td></td>
<td></td>
<td>1. Concern for students' development</td>
<td></td>
</tr>
<tr>
<td>2 Concern for students' development</td>
<td></td>
<td></td>
<td>2. Commitment to scholarship</td>
<td></td>
</tr>
<tr>
<td>3 Commitment to scholarship</td>
<td></td>
<td></td>
<td>3. Working with colleagues</td>
<td></td>
</tr>
<tr>
<td>4 Working with colleagues</td>
<td></td>
<td></td>
<td>4. Equal opportunities</td>
<td></td>
</tr>
<tr>
<td>5 Equal opportunities</td>
<td></td>
<td></td>
<td>5. Reflection</td>
<td></td>
</tr>
</tbody>
</table>

Total = not more than 30 pages
Outcome 5 Evidence Organizer

Use this form to log the evidence you may include for Outcome 5 and for the related values. (Values which do not apply are greyed out.)

Make and use further copies of this form if you find it helpful.

Outcome 5 – Keep appropriate records of own teaching support and academic administrative work.

<table>
<thead>
<tr>
<th>Name or description of item</th>
<th>Date</th>
<th>Source/author</th>
<th>Relevant professional values</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>1. How students learn</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2. Concern for students' development</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>3. Commitment to scholarship</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>4. Working with colleagues</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>5. Equal opportunities</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>6. Reflection</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>No. of pages (max. = 20 overall)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total = not more than 20 pages</td>
</tr>
</tbody>
</table>
**Outcome 6 Evidence Organizer**

Use this form to log the evidence you may include for Outcome 6 and for the related values. (Values which do not apply are greyed out.)

Make and use further copies of this form if you find it helpful.

Outcome 6 – Develop personal and professional coping strategies appropriate to the constraints and opportunities of the institutional setting, to manage adequately time and to operate successfully within available resources.

<table>
<thead>
<tr>
<th>Name or description of item</th>
<th>Date</th>
<th>Source/author</th>
<th>Relevant professional values</th>
<th>No. of pages (max. = 20 overall)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>1. How students learn</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2. Concern for students’ development</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>3. Commitment to scholarship</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>4. Working with colleagues</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>5. Equal opportunities</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>6. Reflection</td>
<td></td>
</tr>
</tbody>
</table>

Total = not more than 20 pages
Outcome 7 Evidence Organizer

Use this form to log the evidence you may include for Outcome 7 and for the related values.

Make and use further copies of this form if you find it helpful.

Outcome 7 — Reflect on own personal and professional practice and development, assess your future development needs and make a plan for own continuing professional development.

<table>
<thead>
<tr>
<th>Name or description of item</th>
<th>Date</th>
<th>Source/author</th>
<th>Relevant professional values</th>
<th>No. of pages (max =20 overall)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Total = not more than 20 pages
# H851/HH851 Teaching in Higher Education

## Portfolio Assessment Form

**Participant's name:**  

**Personal identifier:**  

**Assessor details:**  

### 0 Introduction

<table>
<thead>
<tr>
<th>Requirement</th>
<th>Yes</th>
<th>Overlength</th>
</tr>
</thead>
<tbody>
<tr>
<td>0a Up to 300 word cv</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0b Up to 300 word account of teaching work undertaken</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0c Letter from head of department or course leader authenticating account of teaching work undertaken</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Outcome 1 Design of teaching sessions

<table>
<thead>
<tr>
<th>Requirement</th>
<th>Yes</th>
<th>Overlength</th>
</tr>
</thead>
<tbody>
<tr>
<td>1a Up to 750 word claim</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1b Not more than 30 pp of evidence</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

#### Teaching methods appropriate to:

<table>
<thead>
<tr>
<th>Teaching methods</th>
<th>Clearly achieved</th>
<th>Just achieved</th>
<th>Not quite achieved</th>
<th>Not achieved</th>
</tr>
</thead>
<tbody>
<tr>
<td>1c Learning outcomes</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1d Learners</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1e Mode of study</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1f Subject</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1g Resources</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

#### Values

<table>
<thead>
<tr>
<th>Values</th>
<th>Clearly achieved</th>
<th>Just achieved</th>
<th>Not quite achieved</th>
<th>Not achieved</th>
</tr>
</thead>
<tbody>
<tr>
<td>1h How students learn</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1i Concern for student development</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1j Equal opportunities</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Outcome 1 overall

<table>
<thead>
<tr>
<th>Outcome 1 overall</th>
<th>Outstanding</th>
<th>Clear pass</th>
<th>B‘line pass</th>
<th>B‘line fail</th>
<th>Clear fail</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### Outcome 2 Teaching and learning methods

<table>
<thead>
<tr>
<th></th>
<th>Yes</th>
<th>Overlength</th>
</tr>
</thead>
<tbody>
<tr>
<td>2a</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2b</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>First teaching and learning method</th>
<th>Clearly achieved</th>
<th>Just achieved</th>
<th>Not quite achieved</th>
<th>Not achieved</th>
</tr>
</thead>
<tbody>
<tr>
<td>2c Observed and reported</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2d Use of appropriate learning technologies</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Second teaching and learning method</th>
<th>Clearly achieved</th>
<th>Just achieved</th>
<th>Not quite achieved</th>
<th>Not achieved</th>
</tr>
</thead>
<tbody>
<tr>
<td>2e Observed and reported</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2f Use of appropriate learning technologies</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Values</th>
<th>Outstanding</th>
<th>Clear pass</th>
<th>B'line pass</th>
<th>B'line fail</th>
<th>Clear fail</th>
</tr>
</thead>
<tbody>
<tr>
<td>2g How students learn</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2h Concern for student development</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2i Equal opportunities</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2j Reflection</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Outcome 2 overall

<table>
<thead>
<tr>
<th></th>
<th>Outstanding</th>
<th>Clear pass</th>
<th>B'line pass</th>
<th>B'line fail</th>
<th>Clear fail</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Outcome 3 Marking and feedback

<table>
<thead>
<tr>
<th></th>
<th>Yes</th>
<th>Overlength</th>
</tr>
</thead>
<tbody>
<tr>
<td>3a</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3b</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Values</th>
<th>Clearly achieved</th>
<th>Just achieved</th>
<th>Not quite achieved</th>
<th>Not achieved</th>
</tr>
</thead>
<tbody>
<tr>
<td>3c Feedback to students</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3d Grading or marking</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3e How students learn</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3f Concern for student development</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3g Equal opportunities</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3h Reflection</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Outcome 3 overall

<table>
<thead>
<tr>
<th></th>
<th>Outstanding</th>
<th>Clear pass</th>
<th>B'line pass</th>
<th>B'line fail</th>
<th>Clear fail</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### Outcome 4 Monitor and evaluate teaching

<table>
<thead>
<tr>
<th></th>
<th>Yes</th>
<th>Overlength</th>
</tr>
</thead>
<tbody>
<tr>
<td>4a</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>4b</td>
<td>☐</td>
<td>☐</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Clearly achieved</th>
<th>Just achieved</th>
<th>Not quite achieved</th>
<th>Not achieved</th>
</tr>
</thead>
<tbody>
<tr>
<td>4c</td>
<td>Self</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>4d</td>
<td>Peer</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>4e</td>
<td>Student</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>4f</td>
<td>How students learn</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>4g</td>
<td>Concern for student development</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>4h</td>
<td>Scholarship</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>4i</td>
<td>Working with colleagues</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>4j</td>
<td>Equal opportunities</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>4k</td>
<td>Reflection</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Outstanding</th>
<th>Clear pass</th>
<th>B'line pass</th>
<th>B'line fail</th>
<th>Clear fail</th>
</tr>
</thead>
<tbody>
<tr>
<td>Outcome 4 overall</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td></td>
</tr>
</tbody>
</table>

### Outcome 5 Support and administration

<table>
<thead>
<tr>
<th></th>
<th>Yes</th>
<th>Overlength</th>
</tr>
</thead>
<tbody>
<tr>
<td>5a</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>5b</td>
<td>☐</td>
<td>☐</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Clearly achieved</th>
<th>Just achieved</th>
<th>Not quite achieved</th>
<th>Not achieved</th>
</tr>
</thead>
<tbody>
<tr>
<td>5c</td>
<td>Appropriate records kept</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>5d</td>
<td>Reflection</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Outstanding</th>
<th>Clear pass</th>
<th>B'line pass</th>
<th>B'line fail</th>
<th>Clear fail</th>
</tr>
</thead>
<tbody>
<tr>
<td>Outcome 5 overall</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td></td>
</tr>
</tbody>
</table>
### Outcome 6 Coping strategies

<table>
<thead>
<tr>
<th></th>
<th>Yes</th>
<th>Overlength</th>
</tr>
</thead>
<tbody>
<tr>
<td>6a</td>
<td>Up to 500 word claim</td>
<td>☐</td>
</tr>
<tr>
<td>6b</td>
<td>Not more than 20 pp of evidence</td>
<td>☐</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Clearly achieved</th>
<th>Just achieved</th>
<th>Not quite achieved</th>
<th>Not achieved</th>
</tr>
</thead>
<tbody>
<tr>
<td>6c</td>
<td>Time management</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>6d</td>
<td>Operating within resources</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
</tbody>
</table>

**Values**

<table>
<thead>
<tr>
<th></th>
<th>Clearly achieved</th>
<th>Just achieved</th>
<th>Not quite achieved</th>
<th>Not achieved</th>
</tr>
</thead>
<tbody>
<tr>
<td>6e</td>
<td>Working with colleagues</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>6f</td>
<td>Reflection</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Outstanding</th>
<th>Clear pass</th>
<th>B’line pass</th>
<th>B’line fail</th>
<th>Clear fail</th>
</tr>
</thead>
<tbody>
<tr>
<td>Outcome 6 overall</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
</tbody>
</table>

### Outcome 7 Reflection and planning CPD

<table>
<thead>
<tr>
<th></th>
<th>Yes</th>
<th>Overlength</th>
</tr>
</thead>
<tbody>
<tr>
<td>7a</td>
<td>Up to 500 word claim</td>
<td>☐</td>
</tr>
<tr>
<td>7b</td>
<td>Not more than 20 pp of evidence</td>
<td>☐</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Clearly achieved</th>
<th>Just achieved</th>
<th>Not quite achieved</th>
<th>Not achieved</th>
</tr>
</thead>
<tbody>
<tr>
<td>7c</td>
<td>Reflection</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>7d</td>
<td>Needs analysis</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>7e</td>
<td>CPD plan</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
</tbody>
</table>

**Values**

<table>
<thead>
<tr>
<th></th>
<th>Clearly achieved</th>
<th>Just achieved</th>
<th>Not quite achieved</th>
<th>Not achieved</th>
</tr>
</thead>
<tbody>
<tr>
<td>7f</td>
<td>How students learn</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>7g</td>
<td>Concern for student development</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>7h</td>
<td>Scholarship</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>7i</td>
<td>Working with colleagues</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>7j</td>
<td>Equal opportunities</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>7k</td>
<td>Reflection</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Outstanding</th>
<th>Clear pass</th>
<th>B’line pass</th>
<th>B’line fail</th>
<th>Clear fail</th>
</tr>
</thead>
<tbody>
<tr>
<td>Outcome 7 overall</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
<td>☐</td>
</tr>
</tbody>
</table>

### 8 Overall

<table>
<thead>
<tr>
<th></th>
<th>Yes</th>
<th>Overlength</th>
</tr>
</thead>
<tbody>
<tr>
<td>8a</td>
<td>No outcome or value evidenced wholly by simulation</td>
<td>☐</td>
</tr>
<tr>
<td>8b</td>
<td>Less than 20% of evidence older than 3 years</td>
<td>☐</td>
</tr>
</tbody>
</table>

The final decision, which will be computed, is given by Borderline Pass, Clear Pass or Outstanding Pass on each outcome plus 'Present' for 8a and 8b.
Presenting your portfolio
The Open University

H851/HH851 Teaching in Higher Education

Certificate of confirmation of authorship

I confirm that the evidence and claims in the attached portfolio is my own work, or that any materials included as evidence which are not wholly my own work are clearly labelled to indicate their author(s).

Signed

Name

Date
### Outcome 1 Evidence Record

Use this form to log the evidence you have included for Outcome 1 and for the related professional values.

Make and use further copies of this form if necessary.

Outcome 1 – Design teaching sessions from a course outline, document or syllabus. This involves choosing teaching methods appropriate to the group of learners, the mode of study, the subject material, the resources available and the learning outcomes.

<table>
<thead>
<tr>
<th>Name or description of item</th>
<th>Date</th>
<th>Source/author</th>
<th>Relevant professional values</th>
<th>No. of pages (max = 30 overall)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>1. How students learn</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2. Concern for students' development</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>3. Commitment to scholarship</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>4. Working with colleagues</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>5. Equal opportunities</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>6. Reflection</td>
<td></td>
</tr>
</tbody>
</table>

Total = not more than 30 pages
## Outcome 2 Evidence Record

Use this form to log the evidence you have included for Outcome 2 and for the related professional values.

Make and use further copies of this form if necessary.

**Outcome 2** - Use two appropriate teaching and learning methods from: making presentations (e.g. lectures, demonstrations); facilitating group learning (e.g. through seminars, discussion groups, projects); working with individual learners; facilitating practicals or laboratory classes. You must also include evidence of using appropriate learning technologies.

<table>
<thead>
<tr>
<th>Name or description of item</th>
<th>Date</th>
<th>Source/author</th>
<th>Relevant professional values</th>
<th>No. of pages (max. = 30 overall)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Total = not more than 30
Outcome 3 Evidence Record

Use this form to log the evidence you have included for Outcome 3 and for the related professional values.

Make and use further copies of this form if necessary.

Outcome 3 – Mark or grade, and give feedback on, student work.

<table>
<thead>
<tr>
<th>Name or description of item</th>
<th>Date</th>
<th>Source/author</th>
<th>Relevant professional values</th>
<th>No. of pages (max. =30 overall)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>1 How students learn</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2 Concern for students’ development</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>3 Commitment to scholarship</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>4 Working with colleagues</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>5 Equal opportunities</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>6 Reflection</td>
<td></td>
</tr>
</tbody>
</table>

1
2
3
4
5
6
7
8
9

Total = not more than 30 pages
# Outcome 4 Evidence Record

Use this form to log the evidence you have included for Outcome 4 and for the related values.

Make and use further copies of this form if necessary.

Outcome 4 – Monitor and evaluate own teaching, using self, peer and student feedback

<table>
<thead>
<tr>
<th>Name or description of item</th>
<th>Date</th>
<th>Source/author</th>
<th>Relevant professional values</th>
<th>No. of pages (max.=30 overall)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>1 How students learn</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2 Concern for students' development</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>3 Commitment to scholarship</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>4 Working with colleagues</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>5 Equal opportunities</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>6 Reflection</td>
<td></td>
</tr>
</tbody>
</table>

Total = not more than 30 pages
Outcome 5 Evidence Record

Use this form to log the evidence you have included for Outcome 5 and for the related values.

Make and use further copies of this form if necessary.

Outcome 5 – Keep appropriate records of own teaching support and academic administrative work.

<table>
<thead>
<tr>
<th>Name or description of item</th>
<th>Date</th>
<th>Source/author</th>
<th>Relevant professional values</th>
<th>No. of pages (max = 20 overall)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>1 How students learn</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2 Concern for students' development</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>3 Commitment to scholarship</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>4 Working with colleagues</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>5 Equal opportunities</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>6 Reflection</td>
<td></td>
</tr>
</tbody>
</table>

1
2
3
4
5
6
7
8
9

Total = not more than 20 pages
# Outcome 6 Evidence Record

Use this form to log the evidence you have included for Outcome 6 and for the related values.

Make and use further copies of this form if necessary.

Outcome 6 – Develop personal and professional coping strategies appropriate to the constraints and opportunities of the institutional setting, to manage adequately time and to operate successfully within available resources.

<table>
<thead>
<tr>
<th>Name or description of item</th>
<th>Date</th>
<th>Source/author</th>
<th>Relevant professional values</th>
<th>No. of pages (max. = 20 overall)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td></td>
<td>1. How students learn</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td>2. Concern for students' development</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td>3. Commitment to scholarship</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td>4. Working with colleagues</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td>5. Equal opportunities</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
<td>6. Reflection</td>
<td></td>
</tr>
</tbody>
</table>

Total = not more than 20 pages
## Outcome 7 Evidence Record

Use this form to log the evidence you have included for Outcome 7 and for the related values.

Make and use further copies of this form if necessary.

Outcome 7 – Reflect on own personal and professional practice and development, assess your future development needs and make a plan for own continuing professional development.

<table>
<thead>
<tr>
<th>Name or description of item</th>
<th>Date</th>
<th>Source/author</th>
<th>Relevant professional values</th>
<th>No. of pages (max = 20 overall)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>1. How students learn</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2. Concern for students'</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>development</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>3. Commitment to scholarship</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>4. Working with colleagues</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>5. Equal opportunities</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>6. Reflection</td>
<td></td>
</tr>
</tbody>
</table>

Total = not more than 20 pages

A review paper identifying lessons about the management of staff and educational projects and ventures through analysis of papers published in the International Journal for Academic Development.
Managing EDUCATIONAL DEVELOPMENT PROJECTS

Effective Management for Maximum Impact

Edited by Carole Baume, Paul Martin & Mantz Yorke
Learning from educational development projects

David Baume

INTRODUCTION

Many of the chapters in this book describe educational development projects undertaken in UK higher education. Others look more widely, to project work undertaken in Australia and in Singapore. This chapter looks more widely still, at a number of projects undertaken in various countries (including the UK). The projects discussed in this chapter have all been described in the pages of the International Journal for Academic Development (IJAD). IJAD is the journal of the International Consortium for Educational Development (ICED). ICED was established in 1995 to link together national educational networks and to encourage and support the growth of new networks. More details about IJAD and about ICED and its member national networks can be found on their respective Web sites (see http://www.queensu.ca/idc/ijad/ and http://www.edu.yorku.ca/progers/iced/).

Few of the papers discussing these educational development projects give much explicit attention to issues of project management, being much more concerned with academic and conceptual issues, with data and its understanding and with emergent theory. This relatively low attention given to project management issues is in itself interesting. A thesis of this book is that explicit and thoughtful attention to the management of an educational development project is necessary for the success of the project. However, a discovery from reading journal papers through the lens of project management is that project management has not generally been seen as worthy of explicit academic analysis by educational developers.

Although it is possible to gain, from reading the papers, a wealth of information about approaches to the management of educational development
projects, it has been necessary – as well as fascinating – to dig for the infor­
mation. An additional benefit from this chapter may be that it will help
readers to find information and ideas about the management of educational
development projects, even in accounts of such projects that do not give
much explicit attention to project management.

The bulk of this chapter works from paper to paper, describing for each
one the major issue or issues in project management, drawing implications
for practice from each paper. These implications are considered further in
a concluding section.

Fewer than one-tenth of the papers published in IJAD are considered in
this chapter. Four main criteria were adopted in the selection of papers for
the production of case studies here:

• the accessibility, directly or by clear inference, of information on project
management issues in the paper;
• a range of countries of origin;
• a range of types of projects;
• a range of scale of projects.

PROJECT MANAGEMENT ISSUES ARISING FROM THE PAPERS

Attending to context

Establishing a new educational development unit would match most devel­
opers' conceptions of a 'project'. Sue Johnston and Di Adams (1996) from
the University of Canberra, Australia, stress the importance, for the success
of the project, of attention to context.

One context to which the authors of the paper refer is the existence in the
great majority of Australian universities of an educational development unit,
albeit differing somewhat in structure and roles. The existence of many other
such units, the associated group of educational development professionals,
and the associated professional groupings, conferences and publications,
together provide a positive context, access to peer support and some (though
not total) shared meaning of 'educational development' and 'educational
development unit'.

Implication: Find and use your peer support group(s).

A second context is the particular agenda that the unit was set up to serve, in
this case a growing national concern with quality assurance for teaching and
learning and, fortunately, a related (and at the time well-funded) concern
Learning from Educational Development Projects

for quality enhancement. (The authors comment that 'In many respects, the quality agenda has provided fertile ground for the work of educational development units' (Johnston and Adams, 1996: 20). By similar token the requirement that each English university should produce a Learning and Teaching Strategy has stimulated much development work in English universities.)

**Implication:** Identify and align with national priorities.

A third context is the particular institution in which the unit was to be established, and its particular concerns and values, in this case 'a relatively small Australian university by Australian standards with some 9,000 students' committed in the words of its mission to 'educating professionals, professionally' (Johnston and Adams, 1996: 21). The new unit was to be a vehicle for achieving this.

**Implication:** Identify and align with institutional priorities.

A fourth context was the perception by the staff of the unit, informed by their reading of the institution, of the 'need for the unit to prove itself of worth to the university very quickly and effectively' (Johnston and Adams, 1996: 21) through a full programme of events, publicity, visits, services, projects and contributions to policy.

**Implication:** Anticipate and overcome in advance possible institutional concerns.

There are other valuable lessons for the management of a project in this paper, including: the need, not just for the design of any educational development project to play close attention to the many different contexts in which it needs to function, but also to ensure that its plans and methods give weight to the many, varied and sometimes conflicting needs of those contexts; the need for 'quick wins', for an educational development project to prove itself quickly; the need for political skills.

**Attending to disciplinary demands**

Alan Jenkins (1996) from Oxford Brookes University, UK, describes his 'academic journey from a long period of teaching and researching geography to recently taking on the role of an educational developer'. To summarize his very thoughtful paper, he describes his strongly held view of the central importance of discipline to the practice of academic development. At the start of the paper he quotes with approval Lee Schulman's
observation that 'the key to understanding the knowledge base of teaching lies at the intersection of content and pedagogy', and draws many implications from this powerful statement for the practice of academic development. (I cheerfully acknowledge the profound and lasting effect that this paper by Alan Jenkins has had on my own thinking and practice as an academic developer.)

Implication: The paper has a clear message – academic development work and projects must give substantial and sustained attention to the issues of the teaching of the discipline. Generic approaches by contrast may be less effective.

There is a need for further work, conceptual and experimental, to refine this view, and to identify when and how educational development projects can address generic teaching and learning issues as well as subject-specific issues. But in the meanwhile, Alan Jenkins' plea for attention to the discipline remains good advice for any educational development project.

Sensitivity

Deniz Gokcora (1996) from Portland State University, Oregon, United States, describes a study into a potentially fraught topic. Teaching Assistants (TAs) from the People's Republic of China 'constitute 30 percent of the TA population in the US' (Gokcora, 1996: 34). The study investigated the perceptions held by Chinese TAs and American undergraduates of what makes a good teacher. The differences identified were related to the different views of education held by the two groups – 'Chinese value content, Americans tend to value presentation'. The main method used in the study was focus groups, of TAs (at various stages of progression through a programme on teaching) and of American undergraduates.

Implication: I learn from this paper that, with appropriate respect for cultural difference and with a sensitive research approach, projects can safely and productively investigate and draw implications for action in what may at first hearing appear to be very difficult and sensitive areas.

Holding dreams, attending to reality

Peggy Nightingale (1996) describes the externally funded design and clearly successful introduction of a programme of three courses – graduate certificate, graduate diploma and Masters – on university teaching at the University of New South Wales, Australia. Peggy Nightingale then comments: 'The only problem with trying to be inspirational is that various
realities intrude - the reality of operating within a large bureaucratic system with resource constraints, and the reality that both teachers and students are human.' She then goes on to say 'Six years of experience have driven home some uncomfortable learning experiences for the teaching team.'

The experienced project manager may sigh on reading this; the new project manager may become fearful. They may wonder what particular realities have intruded here, what might have been done in advance to mitigate them or afterwards to overcome them. They may also wonder which of these or similar realities may be about to intrude into the reader's own project. Peggy Nightingale's honest and thoughtful, if somewhat depressing, account is very valuable.

The first realities are the cost and both the academic and administrative workload involved in running a Masters degree in an academic centre that also has many other duties. The administrative workload was all the greater because the centre had to create the infrastructure for running courses.

**Implications:** Development projects by their nature tend to do new and different things. Developers by temperament tend to be optimists, concerned to innovate, to dream and implement those dreams. They may also, perhaps, be unable or unwilling to see the scepticism of others. It may well be that no one has done exactly what your new development project will do; but the odds are that someone has done something in some ways similar. At the planning stage, identify development projects that have some similarities with yours, ask to see the project plans and reports, talk to the project director or manager and ask him or her what he/she did, how well it worked, and what he/she wishes had been done differently. The managers and staff of development projects are generally keen to share experience.

The second realities is the need to compromise with ideals. University regulations about courses, and assessment and completion schedules, were not compatible with the hopes of the course team to run a strongly participant-centred programme. However, the course team, with strong support from the Higher Degree Committee, was sometimes able to negotiate compromises. And sometimes university regulations and requirements were found, in the light of experience, to be helpful rather than the hindrance they were at first thought to be.

**Implications:** Seek early clarity about the framework of rules and procedures within which the project must operate. Fit in with these as far as you can. If any of the rules or procedures really will seriously distort or
damage the work of the project, then, perhaps through one of the project's champions, seek to negotiate changes to or exemptions from those rules or procedures. (Going somewhat beyond the paper, you can sometimes negotiate exemptions from current procedures on the basis that you will conduct a pilot or trial of a modified procedure, and of course evaluate and report back.)

The third realities are the 'uncomfortably complex and ambiguous role relationships' (Nightingale, 1996: 47) between the participants on the programme (academic staff who are subject teachers) and the teachers on the programme (academic staff who are academic developers). To take just one stark example, programme participants are peers of those who deliver the programme, and at the same time are subject to their judgement, to being assessed by them, perhaps with implications for their tenure and progression, and certainly with implications for their self-esteem.

**Implication:** Multiple role-relations such as those that Peggy Nightingale describes are an inevitable feature of life in a complex organization. Those who are managing projects need to anticipate such situations. For example, a project manager may in some settings need to direct the work of a senior academic. Such 'complex and ambiguous role relationships' need to be acknowledged openly, and reviewed and discussed. This may defuse some difficulties before they arise.

**Working on a large scale with several supporting factors**

Many educational development projects use, and indeed rely on, dissemination to enthusiasts and volunteers rather than to the entire potential population of users. Marie-Louise Schreurs and colleagues at Maastricht University, The Netherlands, devised a staff development programme for all experienced teaching staff in the Faculty of Health Sciences at the university (Schreurs, Robertson and Bouhuijs, 1999). The subject of the programme was teaching through problem-based learning (PBL), the predominant teaching method in the faculty.

The staff developers who devised and ran the programme felt that 'a necessary condition for the successful implementation of a large-scale faculty development programme like this is a supportive attitude of the faculty board' (Schreurs et al, 1999: 117). In fact several factors were acting in support of the initiative: external funding, some of which was used to 'buy out' teaching staff time to allow them to attend the programme; the active support of the Dean; and the knowledge that consideration would be given to participation in training when later decisions were made about tenure and promotion.
The project team took many steps to ensure participation. A project team was formed including key senior faculty staff, reporting to the faculty board. To respond to staff reservations about the programme, information about the programme was widely shared, and the programme was emphasized as providing chances to share experience, rather than as being remedial in intent. The Dean chaired an initial meeting of the programme, and allowed time for questions and discussion. Programme topics were chosen after an analysis of the skill requirements for teaching through PBL. Programme workshops were explicitly informal, and based on an explicit model of change process. Participants were encouraged to choose workshops that addressed their particular current concerns. During the workshops, participants produced and planned teaching materials for their future use, rather than just talking about issues and practice. Each workshop was evaluated. Over the two years of the project, 60 per cent of the 180 or so staff eligible took part in at least one event. The authors compare this figure with data showing participation elsewhere in faculty development by only some 5 per cent of tenured staff each year, a comparison they find encouraging to their project.

Implications: Attaining a high degree of participation in a development event or process requires active consideration of each of the many factors that can encourage or discourage participation. In planning a project with ambitious goals for dissemination and reach, it is worth making an effort to identify at least the major success factors, and then planning to maximize each of these factors. Particular success factors may include:

- adequate funding, for the project and for buying out staff time;
- ensuring support at all levels through consultation and continued attention;
- clear needs analysis;
- for a training course, ensuring the immediate practical applicability by participants both of the content of the training, and of the development work undertaken by participants during the training; and
- continuing evaluation and consequent changes to process.

Grounding in theory

Every educational development project has a basis of theory or belief. This basis may not be articulated, but it is always there. Examples are: "This goal (the goal of the project) is attainable"; "The goal of the project is worth attaining"; "The intended methods will successfully and cost-effectively lead
Managing Educational Development Projects

to the attainment of the goal.' Staff development projects, too, embody theories, about what will prompt and support academics to review and change their practices. However, it is unusual to develop a staff development programme explicitly on the basis of a theory about staff learning. Angela Ho (2000) from Hong Kong Polytechnic University describes the development of such a programme.

Angela Ho starts with the recognition that 'university teachers hold personal conceptions that are related to their teaching practices and also to the learning of their students. This has led to the recognition that genuine improvement in teachers has to begin with a change in their thinking about teaching and learning itself' (Ho, 2000: 31). From this starting point, and from detailed consideration of the literature on how teachers change their conceptions, she designed a staff development programme which aimed to bring to the surface, first, teachers' current conceptions about teaching and learning (their espoused theories), and second, the conceptions that were implicit in their practice (their theories in action). The programme aimed then to identify any inconsistencies between their two sets of views about teaching and learning, to offer different theories about teaching and learning, to help teachers formulate revised conceptions about teaching and learning, to work out implications of these for their practice as teachers, and finally to help teachers to commit to adopting appropriate new approaches.

Evaluation of the programme showed: changes in the conceptions of teaching held by a majority of participants; changes in practice made only by those whose conceptions of teaching changed; and, for half the teachers who changed their conceptions, positive effects on the approaches to learning adopted by students.

Implication: An explicit theory or theoretical orientation can form a sound basis for an educational development project and also for evaluation of that project.

Negotiating goals and process

Barbara Grant of the University of Auckland, New Zealand and Sally Knowles of Murdoch University, Western Australia were concerned to help academic women to become academic writers. They saw some of their concerns as practical: helping colleagues to start to write, to increase productivity and satisfaction in writing, and to get published. Behind these, they saw deeper concerns: for example, about how academic women do and do not understand themselves to be writers. They describe the long route, familiar to many people who devise an educational development project, from a broad area of interest or concern, to detailed
analysis of need, to appropriate practical action and its evaluation (Grant and Knowles, 2000).

Barbara Grant and Sally Knowles’s project was a five-day live-in writing retreat. There was an explicit and non-negotiable outcome from the event: the production of an article for publication, or another chapter for a PhD thesis, as appropriate. This was to be a writing event, not just an event about writing. Accepting this outcome, the participants negotiated a programme: one seminar each day on an agreed topic (after lunch), two work in progress seminars (each evening), and lots of quiet writing time. Evaluations were very positive; most resultant papers were completed, and some published; the retreats have been repeated; the community of participants has expanded.

Implications: The ‘project’ of the writing retreat sprang from the authors’ consideration – personal and academic – of the issues of women as academic writers over a number of years. The overall goal of the event was clear and non-negotiable. This goal was also both attractive and defensible to the participants, who had to commit to the project and who had to justify their expenditure of time and money. Within the non-negotiable overall goal, the process itself was negotiated, within a fixed frame of time and venue. The project thus had a solid academic and personal grounding, a clear and attractive goal, and a negotiated process. These qualities will contribute to the success of many an educational development project.

Consulting to achieve embedding

Janice Smith from the University of North London, England and Martin Oliver from University College London, England (Smith and Oliver, 2000) describe a three-year, nationally funded project, EFFECTS (for Effective Framework For Embedding Communications and Information Technology using Targeted Support). The aim was to develop and accredit the capabilities of academics who use learning technologies in their teaching.

There are many possible ways to achieve this. Rather than starting with the design of training courses and materials, the project started at the end, with an account of the seven generic learning outcomes of an EFFECTS programme. (Briefly, these outcomes are: reviewing the use of learning technologies; selecting appropriate learning technologies; planning the integration into practice of appropriate learning technologies; doing so strategically; evaluating impact; disseminating findings; and undertaking one’s own continuing professional development – all underpinned by an understanding of the underlying educational processes.)

This account was developed and refined through a long process of consultation with higher education institutions. Institutions were then free
to develop a course or programme to meet their particular institutional needs, and receive national accreditation for it. (This is very similar to the approach implemented in 1992 by the UK Staff and Educational Development Association (SEDA) to the accreditation of programmes for the accreditation of teachers in higher education (see Baume and Baume, 1996), and the SEDA Web site www.seda.ac.uk. The EFFECTS framework is formally recognized by SEDA.)

Implications: Successful embedding of an innovation requires ownership by those adopting the innovation. First, EFFECTS has shown that long and thorough consultation is an effective way to generate such ownership by the user community. Second, EFFECTS has shown that a simple and clear framework, in this case a qualification framework, defines and helps the achievement of a common standard, while freeing those involved to meet particular local requirements. A third implication of EFFECTS for successful embedding is that the project needs to be clear about what should be tightly specified and what should be left to local interpretation.

Negotiation and consultation; clarity over outcomes and principles; freedom as to local interpretation and implementation. These are three good guidelines for developers working in higher education. Academics value their freedom!

CONCLUSIONS AND SUGGESTIONS

The following overall conclusions and suggestions for the planning and management of staff and educational development projects arose from my reading of the papers referred to. Additional conclusions may of course be drawn. *IJAD* proved a useful source of ideas about the management of staff and educational development projects. The methodology of this chapter – the mining of published papers for insights into project management – can be applied both to other *IJAD* papers and, of course, more widely. These insights will need to be tested for their applicability in your own project and setting.

Context

All educational development projects take place within a particular context and culture. Most take place within many: within a local institution, department or school, programme or course, each with its own policies and priorities; within a particular national higher education system, each again
with its own concerns; and in one or more disciplinary or professional contexts. It is valuable in the early stages of planning a project to identify, explicitly, the main relevant contexts, and then to explore how features of each of these contexts can provide both sources of support to be exploited, and obstacles to be anticipated and mitigated or overcome. Some adaptation and compromise from original project goals and methods may well be appropriate.

**Discipline**

The academic discipline(s) with which the project is to work is a sufficiently important contextual factor to deserve particular attention. There are limits, as yet not well understood, to the effectiveness that generic educational development projects can have on practice within the disciplines, and thus to what can be achieved by educational developers working in an a-disciplinary way. If the outcomes of a project are intended to have effect within a defined discipline(s), then members of that discipline should be involved in planning, undertaking and disseminating the results of the project. If the outcomes are intended to be applicable across all disciplines, then staff from at least a sample of disciplines should be involved.

**Ownership**

Academics are much more likely to adopt, adapt and implement an innovation in their teaching and learning work if they have had some stake in its creation. The widest feasible consultation, on the original project specification and then on successive iterations of the project, develops this ownership and hence increases the chance of adoption.

**Success factors**

What would really make the project fly? What could damage or ground it? Identify the answers to these questions near the start of project planning. Return to the questions at each stage of project review, and be prepared to see new or revised answers. Check whether the success factors are still being delivered and the failure factors avoided, and act accordingly.

**Change plans and goals**

Adapt to changing circumstances. It is a poor idea to stick with a plan that is no longer working properly. It is equally a poor idea to stick with goals that, as a result of changed circumstances, are no longer appropriate or
optimal. It is usually advisable to negotiate such changes of plan or goal with the project sponsor.

Relationships

Members of a project team may well have more than one professional relationship with each other and with project participants and clients. The most obvious of these might be peer, manager and assessor-evaluator. Explicit attention to and management of such different relationships can reduce the scope for friction and discomfort.

Peers and networks

There is fast growing experience in the planning and operation of educational development projects. Conversations with people who have run other projects will identify their learning, and will thus help you to make fewer, or at any rate different, mistakes. Most educational developers are reasonably friendly and approachable – it’s the only way they survive. Connect with local and national networks of developers and project staff.

Framework

Educational development projects are often undertaken in a somewhat atheoretical way, which is understandable given the relatively low (but fortunately fast growing) theoretical base available for educational development. Theories and models from related disciplines – such as education, management, psychology, and organizational development – can offer scaffolding on which to build projects. The use of an explicit basis in one or more theories or models can aid the development and analysis of project plans, and the evaluation of project outcomes. Hopefully you will develop new or modified theories and models, perhaps in response to criticisms of your explicit theories and models.

Rules

There are always rules. The rules of the institution have probably not been designed primarily, if at all, to facilitate the process of development and innovation. Work with as many of them as you can; negotiate ways around those that cause particular difficulties; seek to change only those you fear will be fatal to your project. Find allies who can help with the last two processes.
NOTE AND ACKNOWLEDGEMENTS

No further contact has been made during the preparation of this chapter with the authors of the papers analysed for project management issues. Sole responsibility for the identification and discussion of project management issues lies with me as the author of the chapter. Further, no updating of the content of papers has been undertaken, although the first papers considered here were published in 1996.

I gratefully thank the authors of the papers published in the *International Journal for Academic Development* and used in this chapter; all the other *IJAD* authors; and the referees who have contributed to improving and assuring the quality of papers in the journal. I also thank my co-founding-editors of the *International Journal for Academic Development*, Dr Chris Knapper in Canada and Dr Patricia Weeks in Australia, as well as the new co-editor Dr Angela Brew, for the conversations and shared work which we hope is contributing to establishing and testing a scholarly base for our work as staff and educational developers.

REFERENCES


Publication 5


Within the twin frameworks of outcomes-based evaluation and of three purposes of evaluation – for account, for improvement and for understanding – are proposed an eclectic variety of evaluation approaches for staff and educational development ventures.
Monitoring and evaluating staff and educational development

David Baume

INTRODUCTION

Terminology

This chapter is intended to support the monitoring and evaluation of any activity involving staff or educational development. 'Activity' is intended to serve as a wholly inclusive term, covering a workshop, a programme, a single intervention, a small, medium or large project, or the whole of the work of a staff or educational development unit.

Focus

I devote most of the chapter to evaluation. However, much of what I say about evaluation can also be applied to the monitoring.

Methodology and utility

This chapter begins with a broadly positivist approach, a concern with goals met and deliverables delivered as well as with processes satisfactory. I offer other approaches later in the chapter. The chapter progresses from basic through more sophisticated and varied accounts of the monitoring and evaluation of staff and educational development. It addresses how monitoring and evaluation can lead to increased understanding and to changed (it is to be hoped improved) practice. The later methods described are not necessarily better; fitness for purposes is all. Some ideas from the early methods underpin later methods. However, I believe that the urge for accountability will outlive at least some fashions in monitoring and evaluation. I further believe that, whatever more sophisticated answers and understandings monitoring and evaluation will provide, we developers will
still be asked to show how we have achieved goals, delivered deliverables, and done so through good processes; hence my advocacy of a positivist approach as at least a part of any evaluation.

**Purpose**

My main aims with this chapter are to demystify monitoring and evaluation; to suggest that monitoring and evaluation are an integral part of good practice right from the start of any development activity, rather than optional add-ons; and to offer some particular approaches, with the hope that you will variously review, adopt and adapt these approaches to your own setting.

**Perspective**

I adopt two points of view in this chapter; mainly that of the evaluator, but also sometimes that of the staff or educational development unit or project to be evaluated, which I often call the evaluand.

**Progression**

The early parts of this chapter describe and review a selection of methods in the literature and in current practice. Later in the chapter I propose an alternative account of the purposes of evaluation and offer a new underpinning approach to evaluation.

**SOME BASIC IDEAS IN MONITORING AND EVALUATION**

At its heart, *monitoring* is a very simple and common process. It runs alongside, or indeed is integral to or frequently interspersed with, most conscious human activities. Monitoring involves first asking a question such as 'How is it going?'

Given some answers to this question, monitoring further involves asking and answering a further question along the lines of: 'So (how) should we change what we are doing?'

*Evaluation* is at its heart a similarly simple process, undertaken at the end of or at waypoints during an extended staff or educational activity. Evaluation first involves asking a question such as 'How did that go?'

If the evaluation is being carried out at some intermediate point in the life of an activity, then the answers to this question should inform the answering of further questions along the lines of 'So what should we do next?'
If the evaluation is being undertaken at the end, a different further question will be more appropriate; perhaps 'So what do we learn from this that may be of use or interest to others?'

Thus far, then, I have suggested that both monitoring and evaluation are concerned with goals achieved or not, and with processes satisfactory or not. In all that follows, I encourage you to test the ideas against your own thinking and experience, and also to apply the ideas to a project or unit that you are interested in monitoring or evaluating or having monitored or evaluated.

WHAT WOULD IT MEAN FOR AN ACTIVITY TO BE GOING OR HAVE GONE SUCCESSFULLY OR WELL?

I shall first sharpen the distinction implied in the question above between succeeding and going well. In what follows I shall take succeeding to mean achieving goals, and going well to mean adopting satisfactory or excellent processes. Why is this distinction significant? Consider: An educational development project may have achieved all its goals, but have operated in such a clumsy way as to damage relationships and render further successful collaboration most unlikely. By contrast, a staff development unit may run in a wholly harmonious way, with procedures, budgets and codes of practice properly followed, but with few or none of its goals achieved. Good process, we should hope, is more likely to be accompanied by attainment of goals, and vice versa. But the correlation is well short of 100 per cent.

Having established this distinction, I shall now consider goals and processes, and the rather distinct approaches required to evaluate each, separately.

Goals and their achievement

Goals of three kinds can usefully be distinguished. A staff and educational development activity may specify its goals in terms of what it intends to do; in terms of its activities – for example, 'We shall run twenty workshops.' It may specify what it intends to produce; its outputs – for example, 'We shall produce six booklets and a Web site.' (Clearly, as in the case of workshops, some activities may also be considered as outputs. Sometimes outputs and output-type activities are lumped together as 'deliverables'.) A staff and educational development activity may – one hopes it does – also specify what effects it intends its activities and its outputs to achieve; these effects we may call its outcomes – for example, 'The department will describe the learning outcomes of all its units and will achieve and assure the constructive alignment of learning outcomes, teaching and learning methods, and assessment methods and criteria.'
There is a rather inconvenient relationship among these three kinds of goals. Activities and outputs are easy to describe and quantify, and thus to establish as goals, as suggested in the short examples above. I suggest below further ways to do this. Establishing goals and targets for activities and for outputs makes for easy monitoring and evaluation, and little more than competent management and administration are required to achieve them. But, as developers, we know that we undertake activities and deliver products in order to have effects, on the practices and perhaps also on the knowledge and understanding and world-view of colleagues or whole departments, subjects, institutions. And such outcomes are harder to define and to measure – though never impossible, as I hope the example above suggests and arguments below will confirm.

How might we measure the outcomes of staff and educational development? To the extent that staff development is a form of teaching – and I know that this is a sensitive issue, our clients also being our professional colleagues and deserving of appropriate respect – still, to the extent that staff development is a form of teaching, we can establish intended learning outcomes for our staff development processes, and find out how well participants have attained these intended learning outcomes. This is unexceptionable for programmes for new staff, for example the by now almost ubiquitous Postgraduate Certificate in Learning and Teaching in Higher Education, on which the work of participants is assessed. By contrast, the idea of assessing colleagues at the end of a staff development workshop may produce a sharp intake of breath, from us as we contemplate it and from participants if we try it. My own view is that we all, developers and our clients, need to become a little more robust about this. But what follows does not require this!

So, let us find another approach to assessing the achievement of outcomes. I shall concentrate here on assessing the outcomes of staff development events, not forgetting that the purpose of this exploration is to find ways to evaluate the workshops and, more generally, to evaluate any staff and educational development activities, here to evaluate in terms of extent of goals achieved. This requires a brief excursion into workshop design.

The classic approach to designing a teaching or training event is a linear one. It starts with context, then considers content, participants and aims, then moves through intended learning outcomes to learning and teaching methods and then to learning resources and materials and finally to assessment and evaluation. This classic approach can cause a problem. It can lead to intended outcomes that are, for any number of reasons, unassessable, which also renders the workshop in this important respect unevaluable. The solution to this problem is to iterate in the design, to go back at each stage of design to the previous stage to ensure consistency or, in John Biggs's helpful and mellifluous phrase, constructive alignment.
(Constructive alignment means more than consistency. It also acknowledges learning as a process of constructing rather than absorbing knowledge. I am happy to use Biggs's phrase in its full intended meaning. See Biggs (1999).)

What does consistency mean here? It means that we need to ensure that the workshop outcomes are of a form that can be assessed. This may mean that instead of setting a final test, we can provide a task towards the end of the workshop and invite some or all participants to share their achievements on the task. This is not very difficult, but it does require both sensitivity to the needs and status of our clients and clarity and consistency of thinking on our part. Again, to the extent that staff development is a form of teaching, and further to the extent that we, as developers, will inevitably be judged for what we do as developers and how we do our development as well as for what we say during it; then sensitivity to clients and clarity and consistency of thought are virtues worth practising.

But we may well have outcomes that cannot be assessed in a workshop. For example, we may intend that participants should be able to adopt a new approach to the design of their course, or to create and implement a new assessment strategy or student support system. Such outcomes cannot be achieved in a single workshop. What to do?

We could run a series of workshops, consultations or other processes over time, and at or after the end assess how far our intended outcome for the whole staff development has been achieved, thus informing our evaluation. This long-term approach has much to commend it.

Proxies

There is another approach. We can search for a proxy, an intermediary, some activity on the part of our clients which, while not exactly the intended outcome of our development work, can give us fair confidence that our intended outcome has been achieved.

Chris Rust (1998) has shown how this can be done for staff development workshops. In a careful study, he followed up the effectiveness of 33 workshops delivered by 14 consultants of the Oxford Centre for Staff and Learning Development to over 500 participants. Completing end-of-workshop evaluations, 69 per cent of participants felt that they were very likely to change some aspect of their practice as a result of this workshop and a further 21 per cent considered it possible that they would make such a change. But this was only a statement of intent. What did they actually do?

Four months after the workshop, Rust sent a further questionnaire to those workshop participants who had said that they would be willing to respond to such a questionnaire. Twenty-five per cent of respondents claimed to have changed their practice to a great or a fair extent, and
89 per cent to have made at least some change. A further telephone survey asked a sample of questionnaire respondents about the kinds of changes they had made. Each interviewee provided one or more specific examples of changes they had made as a result of the workshop they had attended. All believed that their changes had been successful. Rust’s conclusions merit reporting in full:

1. Workshops can promote at least some changes in the practice of most participants, and extensive changes in some participants, and these changes [in their practice] can be judged to have been successful.

2. Workshops can successfully reassure the participants and provide them with extra confidence in what they are already doing and confidence to innovate.

3. Workshop ratings are reasonable predictions of likely impact, and indications by participants of how they are likely to change [their practice] are good predictors of likely impact. (Rust, 1998: 72–80)

Rust offers cautions against deducing too much from his findings. The workshops also scored very highly on participants’ satisfaction. The workshops were concerned with practical topics, and changes to practice were intended outcomes. However, even accepting these cautions, these results are surely encouraging for those who plan to facilitate change through workshops. The findings also offer some support to the evaluator who uses participants’ plans to make particular changes as indicators that they will in fact make these changes.

We can draw a broader implication from this result. Where a large project with many similar processes or events is to be evaluated, or an educational development unit is running a number of similar events, Rust’s paper suggests a defensible approach to evaluation. Briefly, this approach is first to identify participants’ stated intentions, and then to follow up a sample to see if their intentions are carried through into practice. This approach could be applied to any monitoring process and to evaluating any development process designed to stimulate changes in behaviour, not just to workshops.

Planning evaluation

I have focused here on staff development activities. Similar principles apply for planning and then evaluating educational development projects, where the demand for evaluation may be even stronger than for staff development. I suggest:
Plan the evaluation as you plan the activity.

Ensure that your evaluation methods will be able to evaluate how far the goals of the project have been achieved.

If you realize that your evaluation methods will not sufficiently evaluate how far the goals of the project have been achieved, consider modifying the goals of the project to make them more readily evaluable.

Is this not suggesting letting the tail of evaluation wag the dog that is the project? Perhaps, but only a little. And I feel that there is something incomplete – perhaps even, though not everyone will agree with me, a little dishonest – about a declared goal, the achievement of which you know in advance that you cannot properly evaluate.

An implication may be drawn from what I have just said that attention to how far the attainment of a goal can be measured invariably leads to a softening of the goal. This is not the case. I have seen an early draft project goal that spoke of 'having a positive impact on the teaching of...'. This could scarcely have been made softer. What it needed was clarifying, sharpening, strengthening. The revised project outcomes spoke of two goals (among others). The first goal was to be the (high) proportion of the teachers in the discipline who would have at least heard about the project, through the proxies of using a number of different and realistic specified methods of reaching them all. These outreach methods were activities or output measures, admittedly, but they provided plausible proxies for modest outcomes. The second goal was to be the (rather smaller) proportion of lecturers in the discipline who would have begun to explore how they might apply some of the project outcomes to their practice, this time through attendance at workshops (activities, yes, but activities that require, or at least very strongly encourage, particular outcomes).

Evaluating processes

I have looked at some length at goals and the evaluation of their achievement. What are possible indicators of a satisfactory or unsatisfactory process? At its most basic, a staff and educational development activity and those who provide it need to comply with legislation and with institutional guidelines, for example on employment and equality of opportunity. Beyond that, it is difficult to provide generally applicable indicators of a satisfactory or unsatisfactory process. Fortunately, it is also unnecessary. We all have instincts for what comprise satisfactory dealings and relationships with others. These instincts get us part of the way there; but probably not the whole way. We also have the ability to talk and listen to colleagues about what they require, expect and aspire to in our dealings with each other. We
may need to apply our considerable expertise in facilitation and negotiation to extract these accounts of requirements, needs and aspirations. We should then probably record them, perhaps in the form of local guidelines and ground rules. (I know that ground rules are a staff development cliché, but they have become so only because they are widely used and often found valuable.) We also need to monitor adherence to them and deal with excursions from them.

Process goals might include the following:

• Everyone is clear about goals and methods.
• Everyone (or a specified subset) has had the opportunity to contribute to project plans.
• Everyone feels that his or her contributions are taken suitably seriously.
• Meeting notes are circulated within 48 hours of the end of the meeting.
• Goals and methods are subject to regular review.

TWO OUTER LOOPS

The evaluability of goals

The discussions above of the need to iterate between various elements of a staff development programme, and about the possible need to modify project goals to make them more evaluable, both embody a broader principle. They both suggest an outer loop of reflection and action, of monitoring and evaluation, as also addressed in Chapter 12. In the case of the workshop, the outer loop checked that the various elements of the workshop were aligned, consistent with each other; that they pulled in the same direction. In the case of the educational development project, the outer loop involved letting the planning of the evaluation inform the setting of the goals of the staff or educational development activity – if you can’t evaluate its attainment, it may not be the most useful goal.

The continued appropriateness of goals

There is a further outer loop; a further-out loop if you will. This further outer loop additionally asks what may sound a heretical question: even if they are evaluable, are the stated goals, the planned processes, still appropriate?

Why do we need to ask such a question? Surely goals and at least some procedures and processes are fixed? Surely the funders will object if we go around changing goals?
The environment in which the activity is being undertaken will surely change, gradually or abruptly or both. Other initiatives will come along that may render the current activity obsolete, or that may duplicate some of the planned goals, or with which some form of co-operation is clearly the only sensible course. For such reasons, we need to check the continued appropriateness of goals. Changes of goal may of course require high-level exploration and approval. But such changes are very likely to become appropriate in the life of a two- or three-year development project, and certainly in the life of an established unit. Better to seek them out than be buffeted by them; better to monitor and evaluate the project or unit in its wider environment than to try to hide from the world.

AN INTERIM CONCLUSION

The suggestions so far would provide a secure basis on which to plan the evaluation of most staff and educational development activities, most projects or development units. The resultant evaluations would be likely to meet most of the needs of those funding the activity or the evaluation.

The second part of this chapter goes beyond, offering some further and richer approaches to evaluation. But if your current needs have been met, this is a good place to stop, for now at least.

A 10-STEP WAY

Staying with the evaluation of goals and processes, but acknowledging more of the complexities that can arise around staff and educational development, a 9-step approach for evaluating staff development was proposed by Baume and Baume (1995), based closely on work by Nevo (1986), here expanded to 10. To the account already given in this chapter, this account particularly adds the identification of stakeholders, their interests and questions, and also the identification of the criteria for the judgement of answers to stakeholders’ questions. It is suggested that substantial additional effort be applied to the planning of evaluation – the first 5 of the 10 steps are about planning. This extended framework also contains a useful wider list of possible types of object to be evaluated.

In brief, the proposed steps are:

1. Identify the object(s) to be evaluated. What you wish to evaluate may be a policy; a development unit or service; a programme; an event or activity; or a project. (It may sound, or even be, glaringly obvious that you should start by deciding what you want to evaluate. It is still worth
taking the few seconds necessary to be sure, as huge amounts of time can be wasted evaluating the wrong thing. And if the few seconds reveal some confusion about what exactly is/are to be evaluated, then the consequent few minutes or hours of further analysis will for the same reason be very useful.)

2. Identify the main stakeholders in the objects(s) to be evaluated. In Weiss's helpful account, stakeholders are 'members of groups that are palpably affected by the [object to be evaluated], and who will therefore conceivably be affected by evaluative conclusions about the [object to be evaluated], or the members of the groups that make decisions about the future of the [object to be evaluated], such as decisions to continue or discontinue funding or to alter modes of operation of the [object to be evaluated]' (1986). If we adapt Weiss's list, the main stakeholder groups are likely to be intended clients and users of the development activity, and their managers; policy makers; staff development unit managers; individual developers; and project staff.

3. Identify the questions or concerns of each major stakeholder or group. These may be about goals, strategies and plans, the approach and activities taken, and the outcomes achieved. A good way to identify their questions or concerns is to ask them what their questions or concerns are.

4. A particularly valuable step is to go beyond stakeholder questions to stakeholder criteria for a satisfactory answer to the questions. These criteria may address four classes of issues:
   - How far are the expressed stakeholders' needs met? (Broadly, these needs should form the goals of the things to be evaluated.)
   - How far are broader institutional or national policy goals achieved or supported?
   - How far are agreed standards, norms and processes met and followed?
   - How effective are the methods compared to other methods that could have been followed?

5. Plan and pilot the methods and instruments to be used.

6. Carry out the evaluation.

7. This is the new step: as well as answering stakeholder questions, also seek to understand the object(s) being evaluated, to make sense of why what was done had the effect that it had. I explore this in more detail below.

8. Report to stakeholders on answers to their questions and concerns.

9. Change staff and educational development practice as appropriate
10. Periodically review evaluation methods and processes.

Baume and Baume (1995) contains a worked example of this method in action.

DEEPER PERSPECTIVES ON EVALUATION

The message 'you are about to be evaluated' rarely lifts the spirits of the impending evaluand. However hard we work to make evaluation an objective, neutral, non-threatening process, being evaluated may well feel like being interrogated, judged and, quite possibly in at least some respects, found wanting. The following approaches will not entirely overcome this evalu-phobia; nothing can. But they variously seek to make evaluation more illuminating, more complete, more appreciative and readier to give due weight to process.

Seeking to illuminate

'It is little exaggeration to assert that educational research has had negligible impact on the workings of educational institutions and on the ways in which academic men and women reflect upon their professional activities' (Miller and Parlett, 1974, preface: 3). Almost 30 years later this is still mostly true, and probably as true for educational evaluation as for educational research—Parlett and Hamilton (1972) propose, and Miller and Parlett (1974) illustrate, a new approach to evaluation. It still feels fresh and highly attractive. The chief aim of the illuminative approach 'is to explore, describe, analyse, elucidate and portray—in other words to illuminate—the practices and processes of teaching and learning, broadly defined, as they occur in their national settings' (Miller and Parlett, 1974: 2). 'The illuminative approach... [is]

- ... problem centred—beginning (as all applied research does) with issues and concerns as defined in real life settings;
- ... practitioner-oriented—designating its chief function to provide information and insight from professional educators;
- ... cross-disciplinary—drawing especially on psychology, sociology, psychiatry and social anthropology for concepts and ways of thinking;
- ... methodologically eclectic—interviews, questionnaires, observation and analysis of documents are used in various combinations, according to the circumstances, defined problems and stages of investigation;
• ... heuristically organised – the research progressively focusing and refining the areas of inquiry as the study unfolds, in the light of accumulating experiences and as the crucial issues-to-be-studied become uncovered' (Miller and Parlett, 1974: 2).

Seeking the whole picture

'Above all, evaluation is the discernment of the good.' Robert Stake (2002), author of this statement, speaks and writes about and practises what he calls responsive evaluation. Stake calls for evaluation to be more holistic, more thoughtful, more experiential. Evaluation, he feels, should find and tell the evaluand's story, should ask 'What's happening here?' He contrasts what he calls responsive evaluation with criterion-based, analytic, information-based evaluation – in fact, with the kind I have advocated in the first part of this chapter.

Stake's evaluations have some of the qualities of a story, giving a rich picture of the setting, the people, the atmosphere, the environment as well as what is happening and why. He advocates a criterion-free response. We cannot, Stake acknowledges, fail to be analytic or use criteria – but we should let analysis emerge, let criteria float to the surface. At the same time, he is happy to approach an evaluation with a question in mind – such as 'Were the goals of the activity achieved?' Stake can be seen as taking forward the illuminative approach described immediately above while maintaining focus on goals achieved if not on criterion met.

Seeking to appreciate

Imagine asking, as a lead evaluation question, 'Think of a time in your entire experience of this [staff or educational development project or unit] when you have felt most excited, most engaged and most alive. What were the forces and factors that made it a great experience? What was it about you, others and your organisation that made it a peak experience for you?' James Ludema and his colleagues (2000), who suggest this evaluation question, call it an unconditional positive question, reflecting the unconditional positive regard that is one of the cornerstones of humanistic psychology. They offer strong rationales for this question as part of an approach that they call appreciative inquiry. I should stress here that Ludema et al are describing a method for inquiry, typically into the functioning of organizations, rather than evaluation. I take responsibility for adapting what they say towards evaluation, although in truth very little adaptation is needed.
First, they suggest, surely one of the goals of an evaluation is to understand and thereby to extend the best, as well as to understand and thereby remediate the worst.

Their second rationale needs a little more space. It hinges on what they see as the quite inappropriate and destructive power of the dominant paradigm, critical inquiry. They quote in Ludema et al (2000), with approval the deliberately militaristic comments of Gergen (1994), who speaks of 'the mammoth arsenal of critic weaponry at our disposal... There is virtually no hypothesis, body of evidence, ideological stance, literary canon, value commitment or logical edifice [or, we might add, staff and educational development activity] that cannot be dismantled, derided or demolished with the implements at hand.' Staff and educational developers are for the most part nice and constructive people, and I am sure that such behaviour is not common – although not all evaluations of staff and educational development are undertaken by developers! Gergen lists some damaging consequences of this relentlessly critical approach, but for our purposes as staff and educational development evaluators or evaluands we should concentrate on the claimed benefits of an appreciative approach, which Ludema and colleagues offer as 'continuously to craft the unconditional positive question that allows the [staff development activity] to discover, amplify and multiply the alignments of strengths in such a way that weaknesses and deficiencies become increasingly irrelevant' (Ludema et al, 2000).

**Seeking good outcomes in good process**

'There is a lot to be said for commenting on the quality of the thinking and processes that [educational development units, organizations and] project teams develop and trusting that good processes tend to evoke good outcomes.' This proposition from Peter Knight (2003) challenged me. I talk above about the importance of monitoring process as well as outcome. But at first it seemed to me to be hopelessly optimistic to trust that good outcomes will follow from good processes.

On the other hand, I note that whether in evaluation or development or many other forms of professional activity, we devote much more attention to making, finding and using good processes, good practices, than to evaluation. This observation in turn suggests that where ultimate goals are impossible to identify, for example within the timescale of the project, and proxy goals cannot be found, then attention to good process may be a plausible proxy for the achievement of good outcomes.
THREE PURPOSES FOR EVALUATION

I have selected above ideas and practices from a large literature about evaluation, and from a much smaller literature about the evaluation of staff and educational development, a range of approaches that I hope will be useful. However, I am conscious that on first (and quite possibly also second) reading they may seem to point in very different directions. Let me offer a synthesis and resolution, in the forms of (in this section) an account of three primary purposes of evaluation and (in the next section) a unifying approach to the evaluation of a staff or educational development activity.

The literature on monitoring and evaluation suggests two kinds of purposes for evaluation: 'formative evaluation' to improve an activity or project as it progresses, and 'summative evaluation' to judge its effectiveness. The distinction has become, I feel, rather dysfunctional. For example, it ignores the often substantial and fruitful overlap in method and data collection of the two forms of evaluation, and thus discourages a more coherent approach to evaluation. It also discourages more fundamental forms of learning from evaluation. I propose instead three possible purposes for the monitoring and evaluation of staff and educational development activities. Like formative and summative evaluation, these three purposes overlap. But their use clarifies, for evaluator and evaluand alike, what a particular evaluation is intended to achieve. These three purposes are to account, to improve and to understand.

**To account**

To account (another term would be audit) means to assure those who funded the project that the project has done and achieved what was said would be done and achieved, and done these things to an appropriate standard and in an appropriate way. The standards and the methods of accounting or auditing must be negotiated and agreed. For example, will 'deliverables delivered' do, or is it necessary to go into the more elaborate processes of operationalizing goals described earlier in the chapter? And what do the two appropriates mean in this particular context? But evaluation as accounting or auditing is a definable and useful function, primarily concerned with satisfying (or of course not) the client, the funder, other stakeholders. Accounting is a part, and may be the whole, of summative evaluation.

**To improve**

'Evaluation can be a form of consultancy and, as such, do a lot for enhancing the thinking and work of those being evaluated' (Knight, 2003). This is a usefully extreme account of what is sometimes called formative evaluation. It suggests the evaluator as critical friend, as someone who is at
Once a part of and apart from the project team, supportive of the broad purposes of the project but all the time looking out for possible inconsistencies in thinking and practice, for mis-steps about to be made or opportunities about to be missed, for productive questions to ask and productive suggestions to make — and often for productive, appreciative silences where no intervention is needed! Evaluating to improve can mean using most of the methods described throughout this chapter, and others as newly discovered or invented as appropriate. Evaluating to improve is not limited to the description that opened this paragraph. But the particular role or roles of evaluator as improver need initial negotiation and periodic renegotiation.

To understand
It is not necessary to understand a staff or educational development activity, in any beyond a superficial way, to account for it or to audit it. By contrast, it is essential to understand what is working and what isn't, and how, and above all why, in order to make confident proposals to improve the activity being evaluated. But, beyond supporting improvement, understanding is surely a valid aim for the evaluation of any staff or educational development activity? Seeking to understand can mean the construction and testing of models and theories and explanations. It can mean employing any of a vast range of disciplinary paradigms, methods, ways of thinking and arguing.

Seeking to understand, almost whatever 'understand' means to the evaluator, evaluand and their clients, is a properly scholarly and academic aim. The negotiation around evaluation to understand will need to address, among other issues, what the parties would find useful forms of understanding and explanation. Claims of increased understanding can be tested in public, like any other academic idea or proposal. If well received, the new understandings can be applied to future projects.

Evaluation as seeking to understand provides a way for academics to be properly academic, to research their practice. I have discussed the crossing from evaluation to research in more detail elsewhere (Baume, 2002). The concept of evaluation as seeking to understand also necessarily puts research back into development projects, from which it is sometimes excluded by funding bodies presumably anxious that funds for development are not side-tracked into research. This phenomenon is described by Wisdom (2002: 128) in the context of a project, History 2000, undertaken under the Higher Education Funding Council's (HEFCE) Fund for the Development of Teaching and Learning (FDTL): 'It is important to remember that the HEFCE had not established a research fund. FDTL was about the implementing of educational change. There would have been little point in finishing History 2000 with the words "Well, that didn't work,

A guide to staff and educational development
90
but at least we know why.” Indeed. The HEFCE would not have been pleased. But how valuable to be able to say, for example, “The project worked in these respects, and here are our explanations for why…”?

NEGOTIATING AND NEGOTIATED EVALUATION

The more usual account of an evaluation boils down to someone (the evaluator) doing something (monitoring and evaluation) to someone or something (the evaluand). Instead, I here characterize the planning of an evaluation, and also much of its undertaking, as a negotiation.

The first outcome of a negotiation about evaluation would be an evaluation plan or contract. This would identify the purpose or purposes of the evaluation, perhaps using the three-part typology of purposes – to account, to improve and to understand – offered in the previous section. It would also describe evaluation methods, reporting, and also the resources to be applied to the evaluation. And it would include the inevitable outer loop, this time a process for reviewing and renegotiating the evaluation process and contract. I concentrate below on some less obvious features of an evaluation agreement:

• An agreement about who will provide and verify which data, and when. For example, the evaluator does not need to attend a workshop to know that it ran, or count heads to know how many attended. More generally, the activities involved in, and the outputs of, staff and educational development are relatively uncontentious. They happened or they didn’t; they were produced or they weren’t. But the evaluator needs to agree how such data about deliverables delivered will be collected and audited, and to what standard of proof. This would support the accounting function.

• Agreement on quality measures or descriptions for the activities and outputs referred to above, or an agreed process for defining these quality measures or descriptions – again supporting accounting.

• An agreed process for sharpening at least some of the project goals to the point where it could be determined to what extent they had, by the end of the project, been achieved. (Another term for this process is SMARTening, which refers to the suggestions that goals should as far as possible be Specific, Measurable, Attractive, Realistic and Timebound.) For example, a project might aim to ‘develop a reflective approach by students to showing how they had attained programme outcomes’. This will need sharpening, or smartening, if its attainment is to be evaluated. This process is concerned with improving, and also with making proper accounting possible.
Agreement on the likely headings for a very short evaluation report very early in the life of a project. This report should focus on the \textit{a priori} appropriateness of the project goals, and also on the appropriateness of project methods and plans in terms of their likelihood of achieving the goals, together with any recommendations for change. This supports improving but also gives a stronger base for accounting.

So, the evaluation plan can be negotiated and agreed. There are subtleties to this. For example, who is the client for the evaluation? Is it the funder of the activity to be evaluated; the activity itself (the staff or educational development project or unit); or some independent agency? The answer to this question affects the power relationship that underpins and informs both the negotiation and the evaluation. When the client for the evaluation is the project itself, and the evaluator’s duties include ‘accounting’ in the sense described above, the evaluator must be willing if necessary to bite the feeding hand.

But what about the evaluation itself as a negotiation? Negotiation implies conversation, exchanges of information and of interpretations, the collaborative development and testing of evolving models and conceptions.

Conversation and negotiation may not seem essential for evaluation for accounting, once the evaluation plan and standards and criteria are agreed. And indeed they are not essential. The evaluator reports, accounts and is finished. But the evaluation is much more likely to be accepted and embraced by the evaluand when the evaluation has been discussed, negotiated and, as far as possible, agreed. And the negotiated evaluation is much less likely to contain errors of fact or interpretation, errors ripe to be pounced on by the evaluand and used to assault the credibility of the evaluation and the evaluator, should the evaluand so wish, for example because the evaluation elsewhere contains well-founded criticisms.

Conversation and negotiation are fundamental to evaluation for improvement. The staff or educational development activity being evaluated for improvement is unlikely to respond warmly to ideas for improvement that fall new-minted from the evaluator’s brain. There are two reasons why: the ideas themselves are likely to be much less good, less grounded, less tested, than if they had evolved in conversation with those involved in the project or unit; and staff and educational developers are not wholly immune to that syndrome that makes academics at once so effective at testing and advancing knowledge and so infuriating to work with – the Not Invented Here syndrome.

And as for negotiation in evaluation to understand – negotiation and debate are simply fundamental to the whole academic enterprise and process, of which evaluating to understand is a part.
Should everything in the evaluation be negotiated? I suggest that it should, for the reasons of quality assurance and acceptability and effectiveness suggested above. Should everything in the evaluation be agreed between evaluator and evaluand? As far as possible – but, ultimately, no, not everything. The evaluator’s terms of reference should be included in the evaluation agreement. These terms of reference should include the forming and expressing of reasoned and evidence-supported independent judgements. Some of these judgements may differ from those of the evaluand. A good evaluation process provides opportunities for the evaluand to comment on the evaluation process and on the reports and judgement the evaluator makes. But the evaluator must evaluate and report.

CONCLUSION

What do I hope you will take from this chapter? Answers to all your questions and solutions to all your problems about evaluation, of course. But, specifically:

• A determination to negotiate and agree the purposes for and the processes of the evaluation.
• A concern to identify stakeholders and the nature of their stakes in the project.
• A concern that evaluation be systematic, which can include being systematically eclectic in methods.
• A concern that evaluation be as far as possible collaborative.
• A concern for the clearest possible project goals as a basis for evaluation; a realization of the limits of the process of clarifying project goals; and some approaches to take when these limits are reached, such as the use of proxy goals or proxy processes.
• Enthusiasm to try a wide range of approaches to monitoring and evaluation.
• A view of evaluation as a scholarly function and a resolution to make it so.
• A determination to build in evaluation, not to bolt it on. Indeed, a determination to make monitoring and evaluation as natural as breathing.
ACKNOWLEDGEMENTS

My thanks for conversations, for sources and indeed sometimes for negotiations about the evaluation of staff and educational development to Jo Tait, Peter Knight, Mantz Yorke, Carole Baume, James Wisdom and members of the various projects to which I have been and am an evaluator.

REFERENCES

Biggs, J (1999) Teaching for Quality Learning at University, SRHE and OUP, Buckingham
Stake, R (2002) Evaluating education, handout from author at workshop at Coventry University, 9 April
FURTHER READING

George, J and Cowan, J (1999) *A Handbook of Techniques for Formative Evaluation*, Kogan Page, London. Quite simply: this book lives up to its title, and will be very useful indeed to anyone contemplating, planning, undertaking and using formative evaluation. A wide and well-chosen range of approaches and instruments are described, and richly illustrated in case studies.

Hamilton, D, Jenkins, D *et al* (eds) (1977) *Beyond the Numbers Game*, Macmillan Education, London. This is a rich and exciting collection of descriptions of, methods for, debates about and reports of educational evaluation. It centres on illuminative approaches.

Wisdom, J (2002) Making an impact: the evaluation of History 2000, in *Managing Educational Development Projects*, ed C Baume, P Martin and M Yorke, pp 125–39, Kogan Page, London. This is a valuable, well-written and thoughtful account of the evaluation of a large educational development project. It pays particular attention to the changing context in which the project takes place.
Publication 6


A later account of the development of teacher accreditation in the UK, focussing on SEDA’s submission and responses to the Dearing Committee and the establishment of the Institute for Learning and Teaching in Higher Education and then on the further development of SEDA’s work around professional development and accreditation.
STAFF AND EDUCATIONAL DEVELOPMENT
CASE STUDIES, EXPERIENCE AND PRACTICE FROM HIGHER EDUCATION

EDITED BY
HELEN EDWARDS, DAVID BAUME & GRAHAM WEBB

CASE STUDIES OF TEACHING IN HIGHER EDUCATION
CHAPTER 17

FAR TOO SUCCESSFUL

Case reporter: David Baume

Issues raised

The issues raised in this case are coping with a success that threatens to damage your organization, and anticipating and providing what your members will need and want in a greatly changed environment.

Background

What is now SEDA, the UK Staff and Educational Development Association, started in the 1980s as an association for two main groups: staff and educational developers, and teaching and support staff who wanted to understand and improve teaching and learning. The organization produced a series of short publications on teaching and learning, and ran annual conferences. In 1992 SEDA launched a scheme to accredit teachers in higher education. The scheme was built on descriptions of the capabilities of higher education teachers, and of the values and principles that should underpin their practice. The scheme did not directly accredit individual teachers. Rather it reviewed programmes to train university teachers. It asked ‘In order to pass this programme, does a lecturer have to demonstrate achievement of the SEDA capabilities underpinned by the SEDA values and principles?’ If the answer was yes and a few more conditions were met then the programme was ‘recognized’, and those successfully completing it were ‘accredited’. Programme recognition and review were undertaken by programme leaders trained for the role. Ten years later, some 2,500 teachers had been accredited and some 65 programmes recognized by SEDA, mainly in the UK but also in Australia, New Zealand, Hong Kong, Singapore and Sri Lanka.
Recognizing and reviewing these programmes generated a significant part of SEDA's activity and turnover. SEDA also developed other accreditation schemes, for example for higher education administrators.

The case reporter was elected chair of the association in 1990, and held the position for five years. He subsequently continued active involvement with the association.

PART 1

I was with the family on a narrow boat on the Brecon and Monmouth canal in South Wales in 1997. It was raining. At the appointed hour I used my shiny, and still rather magical, new mobile phone to ring Alison at the Times Higher Education Supplement.

‘What does Dearing say about teacher accreditation?’ I asked.

‘He's recommended accreditation for all new teaching staff. He says nice things about the SEDA’s work on teacher accreditation. There will be a new Institute to do this accreditation,’ she reported. She read me the relevant sections of the report, then asked ‘How does SEDA feel about all that?’

Rather like the newspaper editor in Chicago, with his two possible headlines already set up - ‘Innocent!’ and ‘Guilty!’ - I had prepared three possible responses. One was for if Dearing said yes to accreditation, one for if he said no, and one for if he said little or nothing about it. I ran through my ‘yes’ responses with Alison - ‘Absolutely delighted... Very important step for the professionalization of teaching in higher education... Glad to see acknowledgment of SEDA’s pioneering role.’

As the call came to an end, I was struck by two conflicting emotions. This must have shown on my face. ‘Well, is it good news or bad?’ asked a concerned Kit, our teenage son. Emotion one - ‘We’ve done it!’ Emotion two - ‘SEDA’s teacher accreditation work produces 30 per cent of SEDA’s turnover. Now what?’

‘Well, it’s a long story, Kit.’ He sighed and returned to the delights of navigating the narrow boat. I ruminated on the news from the phone call.

The year before, in 1996, a National Commission of Inquiry into Higher Education (the Dearing Committee) had been established. It considered, among other things, the training and accreditation of university teachers, and asked interested organizations to make submissions. When the SEDA Teacher Accreditation Committee and Executive met to consider what SEDA should say in its submission to Dearing, the meetings were vigorous and a whole range of views were expressed.

‘Let’s push as hard as we can for teacher accreditation to be national policy,’ said Chris, one of our particularly enthusiastic members.

‘We believe that teaching in higher education should be a professional business and we’ve shown accreditation can work,’ said Carole in her encouraging way.
‘Look, this is a once-in-a-generation chance to try to make a difference – to get teacher accreditation to be national policy,’ enthused Liz.

‘They’ll never go for teacher accreditation – let’s just keep on doing what we’re doing,’ said Geoff, from the modest self-effacing wing of the committee.

It looked like opinions were poles apart, and the submission date was looming.

**Which way will SEDA go in its submission? Why?**

**PART 2**

‘Let’s just go for it.’ This careful and rational analysis won the day at the SEDA Executive meeting. SEDA’s written submission and evidence to the Dearing Committee suggested that all university teachers should undertake accredited development for their teaching work. We said that our experience since 1992 had shown that new teaching staff accepted the need for training in teaching, generally welcomed such provision, and valued a transferable national qualification in higher education teaching. We said that universities valued the clear statements of outcomes and underpinning values that the SEDA scheme offered, as well as the freedom to design and run a programme that best met local needs. We suggested that such programmes might normally comprise a post-graduate certificate. And we also suggested, rather cheekily, that this accreditation should be undertaken through the SEDA Teacher Accreditation Scheme.

‘Students have the right to be taught well’ was the unarguable proposition with which Liz, our chair, and I began our oral evidence to the committee. ‘And training and accrediting teachers can make a powerful contribution to this.’ One or two other organizations also made broadly similar proposals, and the arguments were clearly persuasive. Dearing recommended, among other things, in words I first heard in the Welsh rain, that ‘over the medium term, it should become the normal requirement that all new full-time academic staff with teaching responsibilities are required to achieve at least associate membership of the [Dearing-proposed] Institute for Learning and Teaching in Higher Education, for the successful completion of probation’, and that ‘It should become the norm for all permanent staff with teaching responsibilities to be trained on accredited programmes.’

Members of the SEDA Teacher Accreditation Committee were delighted that our bright idea had been recommended as national policy. We felt that this vindicated the huge effort that had gone into our scheme – from staff developers, course leaders and course participants as well as senior managers who had backed the scheme. More parochially, but understandably, we were delighted that the Dearing approach of national accreditation of courses,
successful completion of which leads to accreditation of individual lecturers, was based on the SEDA Teacher Accreditation Scheme.

But the fear that we had agonized over in committee and that had struck me so forcibly on the boat was now real and immediate — what should SEDA do about its own teacher accreditation work?

What options does SEDA have? What strategies can SEDA adopt?

PART 3

SEDA had little conventional political power. It was not an agency of government or of the universities. It was an independent membership organization, funded by its few hundred members and by sales of publications and conference places. It was big enough to have a full-time professional administrative team, but it was not used either to the limelight or to the corridors of power. We had to learn, and fast.

Well before Dearing reported, it was clear that he was likely to recommend some form of systematic teacher development and accreditation, but it seemed unlikely that he would make precise recommendations on how accreditation should be implemented. There exists in the UK a tradition of a very technical approach to developing vocational qualifications. One such vocational qualification for higher education teaching had already been developed. Finding no support, it had been laid to rest. We were concerned that such an approach might be tried again. We decided, in a further rush of optimism, to see how much further we could push SEDA’s approach to teacher accreditation.

Working closely with the Association of University Teachers (AUT), the trade union that represents staff in the UK’s old universities, we established a working party of the main organizations interested in higher education teacher development and accreditation. To chair the group, we invited a widely respected, recently retired Vice-Chancellor, Clive (now Sir Clive) Booth, from Oxford Brookes University, which was renowned for taking the development of teaching very seriously. With a few phone calls we obtained the necessary funding from the UK funding councils — raising money was never so easy, before or since!

Over a few months, the Booth Committee developed a two-stage qualification for part-time and then full-time teachers, with a progression route between them. The proposed qualification was in many ways similar to SEDA’s, with simply stated outcomes and a set of underpinning professional values. The committee added a new element — required underpinning knowledge, on topics including student learning and different approaches to teaching. The Booth Committee’s proposed scheme fitted well with Dearing’s
recommendations and was generally very well accepted. (Another complex vocational qualification for higher education teachers was proposed at about this time, but again failed to gain significant support.)

In the meanwhile, the Institute for Learning and Teaching in Higher Education (ILTHE) was established and began work in 1999. After wide consultation, it adopted a framework very similar to that which the Booth Committee had proposed. Once again, SEDA was feeling pleased with both strategy and outcome.

Given a target to recruit a large membership very quickly, the ILTHE established, alongside a process to accredit programmes, a direct entry route for experienced staff through submission of an application. It gave blanket accreditation to all current SEDA-recognized programmes and fellowships, and all graduates of SEDA-recognized courses were accepted as meeting the requirements for direct entry to the ILTHE. By the end of October 2002 over 14,000 staff had successfully applied to join the ILTHE.

In the face of this, SEDA's teacher accreditation work seemed very unlikely to have a long-term future. It was a very different environment, with the ILTHE now a very live reality rather than just a plan. SEDA again faced the question of what it should do about its teacher accreditation work.

Given the huge role and scale of ILTHE, what options does SEDA have for its teacher accreditation? What effect will this have on SEDA as an organization?

PART 4

'I'm still going to stay with SEDA at least for the time being even after I've got ILTHE accreditation for my programme.' Tony's remark surprised me. Further probing revealed that Tony, like many programme leaders, knew and liked SEDA's approach. The archetypal account of SEDA's programme recognition and review process was that it was 'both challenging and supportive'. It was 'challenging' because SEDA asked many detailed questions about exactly how the programme met SEDA's requirements, and needed a high standard of proof to be satisfied. It was 'supportive' because SEDA appreciated the good ideas in the programme, shared good ideas from other programmes, and had a relatively informal approach on the day. 'Both challenging and supportive' added up to a shared intention to ensure that each programme was as good as it could be, with few preconceptions about the ways in which a programme can be good.

Some programme leaders stayed with SEDA because they liked the SEDA outcomes and values, and indeed had designed their programmes to meet them. They also wanted to be sure that the ILTHE had a secure future before they abandoned SEDA recognition, despite the costs of gaining and retaining recognition/accreditation from both organizations.
The ILTHE, on the other hand, was keen that SEDA cease its teacher accreditation work on an appropriate timescale, and began a process of negotiation with SEDA. It was concerned about possible confusion in the sector, given two similar teacher accreditation schemes in operation. So SEDA staff found themselves in a quandary.

We saw the logic of stopping our teacher accreditation work – as a development association we had developed a good and effective process, and a close variant of this process was now national policy and fast becoming national practice. On the other hand, we were a membership organization, which could only do, or stop doing, things at the behest of our members. And then there was the money.

After much discussion and debate, we decided to go back to our roots, to concentrate again on staff and educational development. This meant:

- confirming and extending the close working relations we had built with our members;
- shifting our focus from anyone interested in teaching and learning and towards staff and educational developers and others interested to innovate in their teaching and support of learning;
- growing the SEDA Fellowship, the established professional qualification for staff and educational developers, into a development process as well as a qualification;
- acknowledging the fast-changing scene around the enhancement of the quality of teaching and learning in higher education.

The last point included new government-funded agencies and initiatives with very large budgets (tens of millions of pounds) for staff and educational development, providing major resources at low or zero cost to users. It also included the changing demands on staff and educational developers, no longer concerned just with training new teaching staff and running development projects, but now contributing, for example, to the writing and implementation of university learning and teaching and human resources strategies, both bringing substantial government funding.

After several iterations and many consultations we got to the idea of a single SEDA professional development framework – we call this SEDA-PDF. The SEDA-PDF describes a process of professional development and learning, and continues the idea of ‘recognizing’ programmes. It is based heavily on Kolb’s learning cycle and has five elements. Those successfully undertaking and completing a SEDA-PDF recognized programme have:

- identified their own professional development goals, directions or priorities;
- made a plan for their initial and/or continuing professional development;
- undertaken appropriate development activities;
- achieved particular specialist outcomes as described for a named award;
reviewed their development and their practice, and the relations between them.

'Specialist outcomes as described for a named award' needs a little spelling out. Specific named awards (professional qualifications) within SEDA-PDF include 'Supervising Postgraduate Research', 'Supporting Learning' and 'Embedding Learning Technologies'. Each award has a particular set of outcomes as well as the generic development outcomes listed above. For example two specialist outcomes for the Supervising Postgraduate Research named award are the ability (i) to use an appropriate range of methods (and skills) to monitor, examine and assess student progress and attainment, and give feedback on work, and (ii) to supervise production and assessment of the research project (thesis).

In the spirit of SEDA, we don't require that programmes run in exactly this way, only that these processes are clearly visible to and experienced by programme participants. We also retained the use of underpinning values, and additionally adopted Booth and ILTHE's concept of underpinning knowledge. Underpinning and informing the work of those who successfully undertake and complete any SEDA-PDF recognized programme are commitments to:

- an understanding of how people learn;
- scholarship, professionalism and ethical practice;
- working in and developing learning communities;
- working effectively with diversity and promoting inclusivity;
- continued reflection on professional practice;
- the development both of people and of educational processes and systems.

SEDA-PDF embraces both initial and continuing professional development. There are no requirements that a SEDA-PDF programme is at any particular academic level, or of any particular duration, or that it is assessed, or that it brings any particular number of credit points. Explicit attention to the process of development, together with the specified underpinnings, are the core requirements.

SEDA has returned to its roots. How attractive and successful will SEDA-PDF prove to our communities? We don't yet know!

CASE REPORTER'S DISCUSSION

What was it about the approach to accreditation developed by SEDA, and then revised by the Booth Committee, and revised again and implemented by the Institute for Learning and Teaching, that appealed to teachers and staff developers?
From very many conversations with course leaders and course participants at workshops, course validation and review events and assessment boards, I have learnt that people like this approach because they recognize the outcomes as a description of the things they do as teachers. They like the idea that teaching is more than a collection of skills — that it is based on particular values, and underpinned by particular knowledge about learning and teaching as well as about the discipline taught.

Course leaders like the freedom to design their programmes to meet the distinct needs and circumstances of their institution and staff, while holding the assurance of a national standard. The management literature talks of ‘loose—tight coupling’, of the need to specify some things and let people locally determine other things. ‘Tight on outcomes and underpinnings, loose on methods of implementation’ seems to have worked well here.

What else have I learnt over these 12 years? For me this is a story about confidence. We had what seemed to us a good idea, teacher accreditation, and we pushed it out into practice to see whether it would float or sink. I’m sure that another year of development of SEDA’s initial Teacher Accreditation Scheme, and of subsequent schemes, would have made the schemes better. I am also sure that putting them into practice early, and improving them in light of experience and feedback, made them better much faster.

It is a story about how developers work. Academic developers, I feel, must be principled, proactive opportunists. We must be principled in that we know in what broad directions we want to go — in this case, towards improving teaching and learning and towards professionalizing teaching. We are principled also in that we know broadly what values and beliefs inform our practice — in this case, the belief that students have the right to be taught well, and the belief that staff development and accreditation are good ways to help achieve this as well as the values listed above. We are proactive in that we ‘just do’ potentially good things, such as developing and launching a teacher accreditation scheme, and later playing a leading role in setting up the Booth Committee. And we are opportunists in that we must be prepared to hitch our wagon to any star that is passing in a broadly appropriate direction — in this case, to the Dearing Committee and then to the ILTHE.

It is also a story about colleagueship and community. Staff development in UK universities used to be a solitary business — a staff development unit might comprise half a person. The scale of development activity has grown enormously. The existence of a strong and self-supporting association of staff and educational developers made some of the developments here possible. At the same time, the activities supported and strengthened the association and its members. I don’t want to get too Lord of the Rings about this, but a difficult quest and a strong association support each other well.
Acknowledgements

I am enormously grateful to all the many colleagues in SEDA and latterly the ILTHE who have led, undertaken and supported the many and various initiatives described here; to the many programme leaders and staff with whom I have had (often very lively!) debates about the issues and the practices involved in training and accrediting university teachers; to other colleagues across higher education in the other organizations described here for their major contributions to advancing understanding and practice; and to the participants in these programmes who wanted to improve their teaching, and who helped us develop and validate ways to help them do so. This has been a team effort, and will continue to be so, whatever future forms it takes.

Resources

See www.seda.ac.uk

An earlier account of evaluation, which explores then current practice in the evaluation of staff and educational development in the UK, and adapts an existing model to the particular requirements of evaluating development projects.
A Strategy for Evaluation

David Baume and Carole Baume

Before you go I'd like to find out what you thought of the workshop; how it went, what you got out of it, what you think should have been different. If you could just fill in this questionnaire for me...

The feedback questionnaire, or its neighbour the feedback round, is as much a part of the staff developer’s workshop repertoire as the ice-breaker and the goal setting exercise. The obtaining of feedback provides some essential data for the reflective practice of staff development. But how do staff developers plan feedback questionnaires and rounds? How do they decide what to ask? What do they do with the responses? How else is staff development evaluated?

Moving back a step; who are the current audiences for the evaluation? What uses do these various audiences make of the results of the feedback? Moving back still further who should be the audiences for evaluations of staff development? What are their legitimate interests? Against what should evaluations be conducted?

And returning to the vicinity of our starting point for each of these audiences and purposes, what are appropriate methods for gathering, processing, and disseminating the results of evaluations? During 1993 we conducted a survey of practice in the evaluation of staff development in the UK. In summary, this showed that, while evaluation of staff development events and programmes was widespread, much less evaluation was carried out in respect of policy. This finding suggested to us a need to develop a systematic approach to the evaluation of staff development which would embrace policy and strategy as well as methods and processes.

It may be argued, more often, it may simply be felt, that an evaluation system such as the one we outline here is over-elaborate. What problem is this evaluation system designed to solve? In the messy reality that is often the working environment of staff and educational developers, where goals may be unclear or even incompatible, surely all this is a bit idealistic? So why bother?
Directions in Staff Development

• A principled reason: it is surely better to start with an ideal model process, in this case of evaluation, and then where necessary make compromises as a conscious and informed act rather than as a fudge.

• An opportunistic reason: the moves towards increased accountability sweeping the public sector will in due course reach even into corners such as staff development, and staff developers had best be prepared.

• Two political reasons: staff developers can contribute much more to the work of lecturers on the evaluation of their teaching if they have current and first hand experience of evaluation. Further, staff developers’ insistence on the importance of their clients evaluating teaching and learning will be more credible if their own house is seen to be in good order.

• A personal reason: in the wee small hours, or when arguing for a budget, or when making the case for staff development as a respectable profession which makes a real contribution to the quality of educational provision, it is good to be able to go beyond assertion and some way towards proof.

The examples in this chapter are mainly concerned with the evaluation of staff development in direct support of teaching and learning. However, the approach we describe is applicable to all forms of staff development.

Questions about evaluation

We start with some questions about educational evaluation. We then move on to develop a grid or framework within which the various identified stakeholders in the process of staff development can develop, undertake, make use of the results of, and finally review and improve, an appropriate evaluation process.

One difficulty with systems is that they can feel too systematic, dry, mechanical, denying of imagination and serendipity. The system developed in this chapter could indeed lead to dull evaluations; evaluations which miss the life and spark which characterize good evaluation as well as good staff development. A systematic approach to evaluation is needed. The increased pressures for quality, accountability and efficiency all require it. But how do we avoid the lifelessness?

One of the ways we can do that is by remembering that, as well as meeting requirements for accountability, the underlying purpose of evaluation is to understand and improve, in this case staff development and hence the quality of the student learning experience. Evaluation systems and methods must finally be judged in these terms. Human learning, by whoever, remains a complex, fascinating and only partly understood business. In leading to an improved understanding of this learning, the evaluation of staff development should never become dry and dull; it should never miss the surprises from which progress comes.
Questions and initial answers

Nevo (1986) suggests ten dimensions along which evaluation can usefully be considered. He expresses these dimensions as questions, which he then answers in general terms. After listing the questions we modify and extend Nevo's answers to apply to staff development. At the same time we identify the main stakeholders in the evaluation of staff development. The purpose of all this is to establish a framework within which any individual evaluation can be located, and thus to make it possible to devise a valid and locally appropriate evaluation method. Nevo's questions are:

- How is evaluation defined?
- What are the functions of evaluation?
- What are the objects of evaluation?
- What kinds of information should be collected regarding each object?
- What criteria should be used to judge the merit and worth of an evaluation object?
- Who should be served by an evaluation?
- What is the process of doing an evaluation?
- What methods of enquiry should be used in an evaluation?
- Who should do the evaluation?
- By what standards should evaluation be judged?

An evaluation of staff development should comprise a systematic description of the staff development object, followed by a systematic assessment of its merit, worth, value, cost-effectiveness, or other characteristics of interest to the stakeholders.

Three functions can be identified. Formative evaluations are intended to improve the staff development process. Summative evaluations may provide accountability, proving that resources were properly expended; may inform future resourcing decisions; and may inform future decisions on the selection of staff developers and on the form of the staff development process. Evaluation can also serve what Nevo calls a socio-political function, which makes a case for more staff development, or is intended to gain support for particular programmes of staff development. (He reports a fourth function, which he calls administrative: to exercise authority. For most staff development activities this function is probably best considered within the summative function of evaluation.) It is worth stressing that these are not necessarily different types of evaluation; they do not necessarily require different methods of evaluation to be used. Rather they are different uses of evaluation, in some cases making different uses of the same data.

What are the main kinds of staff development objects to be evaluated? We suggest that there are four. A policy for staff development (whether national policy or policy within an institution); a staff development unit or service; a staff development programme; and a staff development event or activity.

Having identified the four broad classes of staff development objects we
may wish to evaluate, the next step is to decide which aspects of these objects should be evaluated. Nevo (1986) suggests: the goals, the strategies and plans, the process of implementation, and the outcomes and impacts. For example, when a university staff development policy is being evaluated, information should be collected on the goals of the policy; the strategies and plans contained within or derived from the policy, which shade into the detailed process of implementation; and finally, the extent to which the policy’s goals were in fact achieved. Additionally, information should always be sought on any unintended effects or consequences.

There are four broad sets of criteria against which any staff development object can be evaluated. First and most directly, the extent to which the immediate expressed needs or goals were met. Then, the contribution which the staff development policy or service or programme or event made to the achievement of broader institutional and even national goals. Easily overlooked in an evaluation, but very important, the extent to which agreed institutional standards and norms in areas such as equal opportunities are met. Finally, the effectiveness of the staff development method adopted compared to other possible methods of achieving the same goals. For example, might a learning package have been more cost-effective than the workshop which was actually run?

The evaluation should meet the needs of each of the stakeholders. Weiss characterizes stakeholders as:

Members of groups that are palpably affected by the programme, and who therefore will conceivably be affected by evaluative conclusions about the programme, or the members of groups that make decisions about the future of the programme, such as decisions to continue or discontinue funding or to alter modes of programme operation.

(Weiss 1986: 187)

Weiss, who is primarily concerned with the evaluation of large American educational programmes, characterizes four major classes of stakeholder: policymakers; those who manage the programmes to be evaluated; those who deliver the programmes to be evaluated; and the clients to whom the programmes are delivered. Translating to the higher education staff development context, a more-or-less standard list of possible stakeholders and their interests can be drawn up for most staff development units, programmes and events and other activities and services in an institution of higher education:

- **Policymakers** may variously be beyond the university, the profession or subject, the government and its various funding, quality and other agencies; and within the university, heads of department/schools, heads of faculty, specified members of senior management, and the university itself and its various committees and boards. Policymakers may want to know what changes should be made to a staff development programme as it operates, but generally only for large programmes. In most cases
policymakers want to know if the programme achieved its goals, and whether the programme should be extended or repeated, or not. Policymakers are interested in summative assessments. They may also be interested in evaluations for socio-political purposes, i.e., to make a case to a further level of decision making to continue funding.

- **Programme managers** are the immediate managers of the staff development service or programme. They need to have the information on which to modify current programmes, to inform future decisions on the use of particular staff developers, and to persuade policymakers and resource managers of the need for continued funding. They need the results of formative, summative and socio-political evaluations.

- **The Practitioners** who facilitate and deliver the staff development need formative evaluations to guide their future practice.

- **The Clients** are immediately the staff who participate, whether they be lecturers, allied staff, course leaders, heads of department or senior management. However, the managers of the participants, and less immediately, but not to be forgotten, the students, are also clients. Participants and participants' managers use evaluations of staff development programmes to inform their future choices about participation. They also gain from the reflection on the staff development process which they undertake in providing feedback.

The main steps in undertaking an evaluation are: clarifying the purpose of the evaluation; planning the evaluation; collecting and analysing the data; and communicating the findings to the various stakeholders. Methods are described in more detail below. Evaluation should be carried out by people with the necessary technical skills for the method to be used – people who understand fully the context, can establish and maintain appropriate relationships and can adopt, adapt or develop a conceptual framework within which the evaluation will be conducted and reported.

The evaluation should be judged above all in terms of its usefulness to the stakeholders. Other criteria include its practicality, accuracy, feasibility (technical and economic) and propriety (with respect to legal and ethical standards).

A system for evaluating staff development

Building on the ideas described above and the survey referred to briefly in the introduction, this section develops a systematic nine-step approach to planning and carrying out the evaluation of staff development, and offers some instruments and some approaches to developing instruments. This process will ensure an appropriate evaluation of any staff development object from an individual workshop or consultation to a major national programme. The list may feel rather heavy for planning the evaluation of a single workshop. A worked example immediately below the list shows how the method
works in practice. We suggest that staff development units and services could use this checklist to devise a set of standard procedures for evaluating each of the main staff development objects of their service, which, continuing to use our adaptation of Nevo's classification, are: policy for staff development, the unit or service, a staff development programme, and an individual activity or event. The nine steps are:

1. **Identify** the staff development objects to be evaluated from the four categories of policy, staff development unit or service, programme, and event or activity. These objects may be a policy, unit, programme or event or plans for any of these.

2. **Identify** the main stakeholders from within the four categories of policy makers, staff development managers, individual deliverers of staff development, and participants and their managers.

3. **Identify** the questions or concerns for each identified stakeholder with respect to four key groups of variable about each staff development object, namely the goals, strategies and plans, process of implementation, and outcomes.

4. Further **identify** the criteria for judging the answers to stakeholders’ questions. These criteria should be developed from four bases: the extent to which the object meets the expressed needs of stakeholders, the extent to which broader institutional and even national goals are achieved, the extent to which agreed standards or norms are met, and the effectiveness of the chosen method compared to other actual or possible methods.

5. **Devise** and pilot evaluation methods and instruments.

6. **Carry out** the evaluation.

7. **Report** to the various stakeholders on their various concerns in an appropriate form.

8. **Make such changes** to current and future staff development practice as are within your area of responsibility.

9. Periodically **review** evaluation methods with respect to their effectiveness and efficiency.

---

**A worked example**

This worked example assumes a relatively conventional allocation of responsibilities between policy maker, service manager and facilitator. Different services with different structures and different management styles will, of course, make for different definitions of responsibility. For example, where the service manager is also the facilitator, responsibilities are combined. However, this chapter is not concerned to pursue issues relating to the management of staff development.

1. The staff development object is a workshop on learning contracts on a new staff course.
2. The main stakeholders in this workshop are the staff development service manager, the facilitator, and the participants and their managers. The policymakers are interested in the new staff programme as a whole, but probably not in an individual workshop.

3/4. Before the workshop, the staff development manager wants to know if the goals of the workshop are consistent with the overall aims of the new staff programme and whether the intended process is consistent with the norms, standards and values of the staff development service. A prior decision will already have been taken in planning the new staff programme that the most effective and efficient way of achieving the overall goals of the programme is through a workshop series of which the current workshop is a part. After the event, the staff development manager wants to know if the goals of the event were achieved. From time to time they will also want to be sure that the norms, standards and values of the staff development service are embedded in individual workshop practice.

After the workshop, the facilitator also wants to know if the goals, and also the intended outcomes, of the event were achieved, and if the methods used were appropriate to the participants and to the norms, standards and values of the service.

The participants want to know whether they achieved the outcomes of the workshop. They want their views to be heard on whether the process was acceptable to them. They want to be able to judge whether the workshop was the best way in which their needs could have been met.

5. A questionnaire format has previously been developed and tested by the service. A version is drawn up which includes the intended outcome of this workshop.

6. At the end of the workshop, the questionnaire is completed by participants and collected before participants leave. The results are collated by the unit administrator.

7. The collated results are copied to the staff development manager, the facilitator and participants. There is no need in this case to produce reports in different formats for different audiences.

8. The feedback reveals a high level of attainment of goals and general satisfaction with the process. A wish is expressed for some selected reading matter to be given to participants a week or so before the workshop to allow more time in the workshop for working up practical ideas. This informs future practice on the new programme.

9. At the end of the first semester of the programme, a few minutes are devoted to reviewing the continued appropriateness of the questionnaire method. Participants say that they want to explore different methods, in particular the use of 'rounds', partly for variety and partly to give experience of different methods which they can use in their own teaching. This is agreed, but the programme leader explains that the forms will also have to continue to be used to establish comparative data for the course from year to year for quality assurance purposes.
The University of Westsea  
Staff Development Service  
Workshop Planning and Evaluation Form  
Course: Programme for new teaching staff  
Event: Workshop on Learning Contracts

<table>
<thead>
<tr>
<th>Intended outcome of the workshop:</th>
<th>Evaluation (circle or underline)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Participants will have started to develop an approach to using learning contracts in one of the units on which they teach.</td>
<td>To what extent did you get each out of the workshop?</td>
</tr>
<tr>
<td>2</td>
<td>1 Fully/Mostly/Partly/A bit/Not</td>
</tr>
<tr>
<td>3</td>
<td>2 Fully/Mostly/Partly/A bit/Not</td>
</tr>
<tr>
<td>4</td>
<td>3 Fully/Mostly/Partly/A bit/Not</td>
</tr>
<tr>
<td>5</td>
<td>4 Fully/Mostly/Partly/A bit/Not</td>
</tr>
<tr>
<td>6</td>
<td>5 Fully/Mostly/Partly/A bit/Not</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Evaluation</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>To what extent was the workshop:</td>
<td></td>
</tr>
<tr>
<td>6 Interesting?</td>
<td>6 Fully/Mostly/Partly/A bit/Not</td>
</tr>
<tr>
<td>7 Informative?</td>
<td>7 Fully/Mostly/Partly/A bit/Not</td>
</tr>
<tr>
<td>8 Enjoyable?</td>
<td>8 Fully/Mostly/Partly/A bit/Not</td>
</tr>
<tr>
<td>9 Conforming to published staff development service norms?</td>
<td>9 Fully/Mostly/Partly/A bit/Not</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Planning future work</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>11 Would future workshops, consultancy, published information, whatever, on this topic be useful? (If so, please suggest what form these might take)</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Name and Department (optional)</th>
<th></th>
</tr>
</thead>
</table>
Undertaking each step of the system

After the worked example, there now follows a more general account of each step of the system, with suggestions on how the system can be adapted to particular situations.

Identifying the staff development objects to be evaluated

The first step is simply to identify the major staff development objects for which you have some responsibility. It is useful to categorize these under appropriate headings chosen from policies, staff development unit or service, programme, and event or activity. It is important to consider as evaluation objects all the rich variety of services provided, including products (leaflets and newsletters), short informal consultations and use of a resource centre.

The second step is to decide on an overall evaluation strategy and time scale within which each of the objects identified will be evaluated. Some of these decisions will already have been made at institution level. For example, there is likely to be a policy on the frequency and conduct of the review of programmes and courses, and perhaps also on the review of units and departments. On the evaluation of policies there may be little or no clear guidance, and it may be up to the manager of a staff development service, for example, to decide on an appropriate frequency for reviewing the unit’s policies and strategies. At the other end of the scale decisions will have to be made about the frequency with which individual events are evaluated. We suggest starting by evaluating everything and then reducing frequency if the costs, in terms of time and even ‘evaluation fatigue’, start to outweigh the value of the data generated. Like everything else, an evaluation strategy needs evaluating!

Identifying the main stakeholders

The grid below, and the subsequent commentary, is intended to help to identify the most likely stakeholders in each type of staff development object:

<table>
<thead>
<tr>
<th></th>
<th>Policy</th>
<th>Unit or Service</th>
<th>Programme</th>
<th>Event</th>
</tr>
</thead>
<tbody>
<tr>
<td>Policymaker</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>Service manager</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
</tr>
<tr>
<td>Facilitator</td>
<td>9</td>
<td>10</td>
<td>11</td>
<td>12</td>
</tr>
<tr>
<td>Participant</td>
<td>13</td>
<td>14</td>
<td>15</td>
<td>16</td>
</tr>
</tbody>
</table>

Policymakers will have a legitimate interest in the evaluation of policy (1), and in the evaluation of units or services which deliver that policy (2). They may also have some interest in overall programmes, for example, those designed to move an institution towards greater use of resource-based learning (3), but they will be less interested in the evaluation of the design and delivery of individual events (4).
As an implementer of some university policies, the manager of a staff development service or unit is normally interested in, and hopefully contributes to the evaluation of, university policy (5). They will be involved in the evaluation of the policies of the unit or service itself. However their main concern is in the evaluation of the unit or service (6) and of the programmes run by their service (7). Their interest in the evaluation of specific units or events may be confined to the issues of attainment of goals and adherence to standards rather than fine evaluation detail (8). However they are likely to have a management and a developmental role for the facilitator who delivers the activity or event.

The facilitator may be asked to contribute to the evaluation of policy at service or unit level (9) and to the evaluation of the unit or service of which they are a member or for which they have worked (10), but their main concern is with the evaluation of individual programmes (11) or events (12) for which they are responsible. Participants in events or programmes are not generally involved in the evaluation of policies (13). They may be asked to contribute to the evaluation of the service (14). They should have a major input into the evaluation of a programme (15) and individual events (16). They will also make evaluative judgements about the unit or service, programme or event, whether or not they are asked to do so!

Identify the questions or concerns, and the criteria for judging the answers, for each identified stakeholder

What follows are suggestions. The only safe way to identify stakeholders’ questions and concerns is to ask them!

Stakeholder: Policymaker

<table>
<thead>
<tr>
<th>Policy</th>
<th>Unit or Service</th>
<th>Programme</th>
<th>Event</th>
</tr>
</thead>
<tbody>
<tr>
<td>Goals</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Strategies and plans</td>
<td>5</td>
<td>6</td>
<td>7</td>
</tr>
<tr>
<td>Process of implementation</td>
<td>9</td>
<td>10</td>
<td>11</td>
</tr>
<tr>
<td>Outcome</td>
<td>13</td>
<td>14</td>
<td>15</td>
</tr>
</tbody>
</table>

Policymakers are likely to be very interested in evaluations of the goals (1), strategies and plans (5), process of implementation (9) and outcomes (13) of policies. They will also be interested in the goals (2) and the outcomes (14) of units and services charged with implementing the policies. Being concerned with policy and the success of its implementation they will probably be little interested in details of unit strategies and plans (6) and processes of implementation (10), and, as previously established, interested much less or not at all in programmes and individual events.

Within this framework, policymakers’ questions on the goals of the policy and the unit or service will be about their appropriateness to wider goals and policies. Their questions on strategies and plans will concern their feasibility. Their major concerns on processes of implementation will
probably be on resourcing. They will be interested in the extent to which planned outcomes are achieved, and in any unexpected outcomes.

How will policymakers judge the answers to the evaluation questions in which they are interested? What criteria will they use? They will be interested in four things: the extent to which the policy, and the unit or service charged with delivering that policy, meets the university's needs; the extent to which the policy contributes to the university's mission and even to national goals; the extent to which the agreed standards or norms of the university are met by the staff development service; and the effectiveness of the policy and the staff development service compared to other possible policies or systems for delivering staff development.

**Stakeholder: Staff development service manager**

<table>
<thead>
<tr>
<th>Policy</th>
<th>Unit or Service</th>
<th>Programme</th>
<th>Event</th>
</tr>
</thead>
<tbody>
<tr>
<td>Goals</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Strategies and plans</td>
<td>5</td>
<td>6</td>
<td>7</td>
</tr>
<tr>
<td>Process of Implementation</td>
<td>9</td>
<td>10</td>
<td>11</td>
</tr>
<tr>
<td>Outcome</td>
<td>13</td>
<td>14</td>
<td>15</td>
</tr>
</tbody>
</table>

Service managers may be invited to make some contribution to evaluating the goals and outcomes of policies (1 and 13), but they will be more involved in the evaluation of strategies and plans for delivering those policies (5 and 9). They will be heavily involved in evaluating all aspects of their service (2, 6, 10 and 14). They will be interested in evaluating all aspects of programmes run by their service (3, 7, 11 and 15), and in the goals and outcomes of particular events (4 and 16), though to a lesser extent in detailed strategies and plans and processes of implementation (8 and 12).

What criteria will service managers use to judge the answers to the evaluation questions in which they are interested? Again they will be interested in four things: the extent to which the service and its activities respond to and meets the needs of the university and its members; the extent to which the service and its activities contribute to the university's mission; the extent to which the agreed standards or norms of the university are met by the staff development service and its programmes; and the effectiveness of the current programme of events compared to other possible methods of delivering staff development.

**Stakeholder: Staff development event facilitator**

<table>
<thead>
<tr>
<th>Policy</th>
<th>Unit or Service</th>
<th>Programme</th>
<th>Event</th>
</tr>
</thead>
<tbody>
<tr>
<td>Goals</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Strategies and plans</td>
<td>5</td>
<td>6</td>
<td>7</td>
</tr>
<tr>
<td>Process of implementation</td>
<td>9</td>
<td>10</td>
<td>11</td>
</tr>
<tr>
<td>Outcome</td>
<td>13</td>
<td>14</td>
<td>15</td>
</tr>
</tbody>
</table>
The facilitators of an individual staff development event are only concerned with the evaluation of their event, albeit they are concerned with all aspects of that event. Where the facilitator is responsible for a whole series or programme of events then their involvement in evaluation, of course, increases accordingly.

Again the facilitator will use four sets of criteria: the extent to which the event (or programme) meets the published goals of the programme (and, if these are sought at the start of the event or programme, the individual goals of the participants); contributions to any broader goals of the service and the university of which the facilitator has been made aware; as with the other stakeholders, the extent to which to the agreed standards and norms of the service and the programme are met within the event or programme; and, finally, the effectiveness of the current methods used within the event or programme compared to other methods which might have been adopted.

### Stakeholder: Participants and their managers

<table>
<thead>
<tr>
<th>Goals</th>
<th>Policy</th>
<th>Unit or Service</th>
<th>Programme</th>
<th>Event</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strategies and plans</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
</tr>
<tr>
<td>Process of implementation</td>
<td>9</td>
<td>10</td>
<td>11</td>
<td>12</td>
</tr>
<tr>
<td>Outcome</td>
<td>13</td>
<td>14</td>
<td>15</td>
<td>16</td>
</tr>
</tbody>
</table>

Participants and their managers are interested in the extent to which the stated goals of the events or programme, and any individual goals they have, are met and the desired outcomes are achieved. They are also interested in the appropriateness and efficiency of the strategies, plans, and methods adopted.

### Devising, piloting and carrying out evaluation methods

The design and conduct of evaluations can be a complex business, and some thorough guides are available (see for example, Cronbach 1982; Morris 1990 and Tessmer 1993). What follows does not replace those. Rather it introduces some of the key issues involved in the practical evaluation of staff development.

If the purpose of the evaluation is formative, then the evaluation clearly needs to be conducted, analysed and considered in time for the results to influence the current operation of the event, programme, service or policy. Timing of evaluation will also depend on the outcomes being evaluated. For example, the attainment of outcomes concerned with changed behaviour back in the workplace cannot be evaluated during a workshop.

The question as to the appropriate sample size is normally answered on grounds of economics and feasibility. It is simple and useful to obtain the views of all six participants in a small workshop, still manageable and useful to gain feedback from all 36 participants in a programme, but certainly not
feasible to conduct a detailed assessment of the views of 606 recipients of a newsletter. Within some overall goal, such as spending 5 per cent of the unit's time on evaluating its activities, detailed allocations of effort can be made and then reviewed in the light of the value of the data collected.

Ideally for summative evaluation, someone not directly involved in the delivery of the programme or the management of the service could usefully conduct the evaluation. However, formative feedback may most usefully be collected by the person carrying out the activity. This would be particularly true, for example, when collecting feedback at the end of the first day of a two-day workshop in a way which can influence the delivery of the second day.

Questionnaires can obtain participants' views on stakeholders' concerns and issues. Guidance on questionnaire design can be found in Oppenheim (1992) and many other sources. Open-ended questionnaires can gather rich data which can be slow and expensive to collate and analyse. The results of closed questionnaires are easy to collate and analyse. A range of intermediate methods is available. For example, a small sample of interviews can be used to generate statements which can then be used with multi-point agree–disagree rating scales in a questionnaire. An appropriate method for the evaluation of policy might be a questionnaire sent to every member of staff affected by the policy on its continued appropriateness, the methods being used for implementing it, and the outcomes of the policy as the respondents experience them. The questionnaire could be preceded or supplemented by a series of semi-structured interviews.

Oral feedback can be conducted through a round at the end of a workshop, or programme. Participants can be asked open questions, for example, ‘What was the best thing for you about this workshop?’; ‘In what one respect should the next workshop be different?’ Oral feedback assumes that the participants are willing to be open with each other and with the facilitator.

If a service or policy is being evaluated, a series of semi-structured interviews may be a more appropriate tool, allowing as it does the evaluator’s concerns to be addressed while allowing space for the interviewee to voice concerns which were not the subject of specific questions.

A staff development service can be evaluated with respect to its various goals, methods and outcomes by a combination of telephone or questionnaire enquiries to users of the service (and to people entitled to use the service who have not done so), and by the collation of the results of evaluations of its programmes and events. Again, interviews might also be undertaken, preferably by an independent evaluator.
extent to which formative evaluations provide clear bases for appropriate action, and the efficiency of the evaluation process. It will generally be the case that summative evaluations are found more useful when they comprise mainly numerical data, while formative evaluations are more useful when expressed in words and provide guidelines for future action. On the basis of evaluation reports, actions can be determined to change the policy, service or practice being evaluated. The details of how to make these changes, how to use the results of the evaluation, take us outside the scope of this chapter.

Reflection

As we suggested in the introduction, all this, though clearly worthy and rigorous, may still feel a bit much. On reflection, it seems to us that to ask just one or two of Nevo's questions of the staff development currently being planned or undertaken is still useful, even if time and other pressures do not allow the full evaluation system to be used. Who are the main stakeholders and what do they want? What criteria should be used to judge the worth of the piece of staff development to be evaluated? Even in isolation, these are powerful questions.

It also seems to us that, like many activities, systematic evaluation becomes rapidly easier with a little practice. What looks initially daunting becomes, after the second or third round, routine, though hopefully still useful. Not all of the questions need to be asked all the time. For example, policies change slowly, and the stakeholders for each of six workshops may be the same.

The idea that all professionals are reflective practitioners is fast becoming a tired cliche. Reflection alone can slide into empty cogitation. Reflection needs evidence on which to reflect, data to process. The system described here provides that evidence and data on which individual staff developers can sharpen up their practice and on which staff development as a profession can grow and demonstrate its rigour and worth.

Acknowledgements

Graham Gibbs and James Wisdom gave us feedback on an earlier draft. We followed most of their advice, and the chapter is much better as a result. We thank them.

An account for a non-specialist audience of the work reported in much more detail in Publication 9 on the reliability of assessment of portfolios and some implications of this work for the improvement of assessment of portfolios and more broadly.
ASSESSMENT
CASE STUDIES, EXPERIENCE AND PRACTICE FROM HIGHER EDUCATION

EDITED BY
PETER SCHWARTZ AND GRAHAM WEBB

CASE STUDIES OF TEACHING IN HIGHER EDUCATION

KOGAN PAGE
Issues raised

This case focuses on the issue of optimizing reliability in the assessment of portfolios.

Background

The UK Open University (OU) is a large, distance-learning university. Established in 1970, it now has some 200,000 students (over 2,000,000 during its lifetime), 1,000 academic staff, 2,000 support staff, and 7,500 part-time tutors (called Associate Lecturers). Most full-time staff are based at Walton Hall in Milton Keynes. The University has 13 regional centres across the United Kingdom. The University’s Centre for Higher Education Practice (CeHEP) was set up in 1997 to develop courses to train university teachers (as well as to conduct research on the training of university teachers and to run major national educational development projects).

The research described in the case was undertaken by the authors while Mantz Yorke was visiting professor in CeHEP. The ‘we’ in the case refers to the course team when describing course planning and operation, and to the authors when describing the research.
PART I

We were in the process of analysing the first two sets of assessment results for The Open University's new course on teaching in higher education. What? Only 60 per cent pass/fail agreement between assessors overall? After all the work we had put into planning our assessment, how could this be?

In developing an assessment strategy for the new course, we had faced these questions:

- How do you assess the capability of professionals, in this case professional lecturers who are also colleagues?
- How do you make that assessment valid, reliable, rigorous, fair, and possessing all the other virtues we'd hope for in a good assessment system?
- And, this being The Open University, how do you plan to do all this for hundreds or thousands of people at a time?

Our answer had been portfolios, with lots of effort applied at every stage of the teaching and assessment processes to maximize reliability.

Why portfolios? Three main reasons.

First, building a portfolio means assembling real evidence, in this case evidence of teaching ability, rather than writing essays or reports about teaching. It means the lecturer saying: 'Here is a picture of me, teaching and assessing students.' Portfolios can mean valid and authentic assessment.

Second, building a portfolio means being reflective and critical about the evidence and what it suggests. It means the lecturer saying: 'Here's what I think about my teaching. Here's why I teach the way I do. Here's why I think the teaching was good, or less good. Here's how I am improving it.' Portfolios require reflection.

Third, building a portfolio additionally means making an explicit claim about teaching ability. In the portfolio the lecturer says, in effect, 'Here's how this evidence and this reflection show that I have attained each of the course outcomes in turn.' This kind of portfolio involves self-assessment.

We also identified two minor reasons for using portfolios. What better way for teachers to learn about portfolio-based assessment for their students than to assemble, and be assessed on, their own portfolios? Also, a portfolio, unlike an exam script, can live and grow over time, recording continuing professional development as well as initial qualification.

The qualifications environment in which we developed the courses on teaching in higher education had a huge effect on the design of the courses and the assessment process. In 1992, the UK-based Staff and Educational Development Association (SEDA) launched a process to accredit the professional ability of those who teach in higher education. The SEDA Teacher
Accreditation Scheme 'recognized' programmes for training higher education teachers. To be recognized by SEDA, a programme must above all ensure that those who successfully complete it have met specified outcomes, and have done so in a way demonstrably underpinned by specified principles and values. In summary, the abilities required are to: plan teaching sessions; teach; assess student work; monitor and evaluate teaching; keep records of teaching; cope with the demands of the job; and reflect on practice and continue professional development. These abilities must demonstrably be underpinned by an understanding of how students learn; a concern for students’ development; a commitment to scholarship; a commitment to work with and learn from colleagues; concern to ensure equality of opportunity; and continued reflection on professional practice.

In 1999, a new professional body, the Institute for Learning and Teaching (ILT), was set up in the United Kingdom for those who teach in higher education. The ILT accredits courses to train higher education teachers and admits into membership those who pass them, as well as allowing entry by direct application to ILT. The ILT specifies five main areas of work in which teachers must show their ability, underpinned by specified professional values and knowledge.

CeHEP developed a programme of three courses, each portfolio-based, accredited variously by SEDA and ILT. This case study concerns the first of these courses, on teaching in higher education, which leads to the SEDA Associate Teacher qualification, and in due course, it is hoped, also to ILT Associate Membership. The team designing the course was concerned to make assessment as valid and as reliable as possible. Some of the steps taken to achieve these two goals were adopted or adapted from normal OU practice; others were developed especially for the course.

The portfolio that each participant produces for the course on teaching in higher education is arranged in sections, one section for each of the seven course outcomes. Each section of the portfolio contains two main kinds of material: evidence and claim.

Evidence may include, for example, lesson plans, assessed student work, student feedback, or reports of observations of teaching. Most of this evidence will have been produced anyway during the teacher’s work.

The claim is an explicit and reasoned assertion by the lecturer that the evidence shows that he or she has achieved each specified course outcome, and has done so in a way that is underpinned by the specified principles and values. The claims are specially written for the portfolio.

Detailed guidance is given on each step of writing the portfolio. Sample claims, with tutor comments, are also provided. We did not want the business of writing portfolios to get in the way of the lecturers thinking about, developing and showing their abilities as teachers. Participants receive detailed written feedback from their tutor on several draft sections of the portfolios.
We cannot assume that all course participants have access to a good library on teaching and learning in higher education, so we provide course materials. These materials are of two kinds: Practice Guides on the major topics in teaching and learning, and Reader Chapters that provide, with commentary, some of the key theory and evidence that underpins the practical guidance.

A very detailed framework for assessment of the portfolio is provided to participants. We want it to be absolutely clear what 'passing the course' means, in terms of abilities developed and demonstrated. A total of 75 assessment judgements are recorded for each portfolio. This detailed prescription is not an attempt to remove judgement from the assessment process; it is an attempt to make the basis on which the judgements are being made as explicit as we can. Each of the seven outcomes must be passed for the course as a whole to be passed. An outcome can be passed as long as there is no more than one marginal fail on one component of that outcome.

To match the detailed information given to participants about the portfolio and its assessment, we give extensive guidance and training to those who will be assessing the portfolios. This includes, in addition to the framework of assessment requirements, detailed instructions on how to assess the portfolios. (Copies of the instructions for assessors are also given to course participants, as we want assessment to be an open process.)

During preparation for each round of assessment, all assessors are given the same one or two current portfolios as samples to read and assess. The assessors then get together for a day, share their judgements, and strive to reach agreement. This process leads to further clarification of the assessment criteria and standards the assessors will use when undertaking real assessments over the next couple of weeks. Advice emerging from this coordination process is also fed forward to participants and assessors on the next presentation of the course.

During the actual assessment, each portfolio is assessed by two assessors – tutors on the course – who assess independently. If they agree, their recommendation goes forward to the assessment board. If the two assessors disagree whether a portfolio should pass, a third assessor, usually a member of the team who produced the course, moderates the assessment of the portfolio. This third assessor's judgement becomes the recommendation to the assessment board.

So – everything is done in accordance with good educational practice and in the best traditions of the OU. Assessment requirements and processes are specified in great detail, and given to everyone involved; assessors are carefully selected, trained and briefed; assessment judgements are coordinated. A recipe for success?

Well, up to a point! As we looked at the mountain of data generated during the assessment of the first batch of 53 portfolios, we were variously pleased and startled at some of what we found. The pass rate at the level of individual outcomes was 92 per cent. There was reasonable agreement
between the two assessors on whether a given candidate had passed a given course outcome – 87 per cent agreement. However, there was a much lower degree of exact agreement on outcomes (we used a five-point scale for assessing outcomes) – only 39 per cent. And – the judgement of most importance to candidates – first and second assessors agreed on overall ‘pass/fail’ for only 60 per cent of the portfolios.

We were disappointed. So we took a hard look at what had happened. We asked ourselves some searching questions: Why, despite all our efforts, was the overall assessment reliability no higher? What could we do to improve the reliability of assessment of the portfolios, immediately and in the medium and longer terms?

What do you think might account for the lower than hoped-for level of inter-assessor reliability?
What steps would you recommend taking to improve the situation?
What do you think actually happened?

PART 2

We first tried to understand why the reliability of our assessment was lower than we had hoped. A trawl through the sparse literature on portfolio assessment showed a spectacular range of levels of agreement between first and second assessor, ranging from 19 to 100 per cent. However, some of these results described a single overall assessment judgement, others an assessment made up from several different elements. As is so often the case, comparison was difficult, but we derived some little cheer from being in the middle of a range of reported degrees of agreement, rather than near the bottom! Our interpretation of the literature did, however, highlight for us the tension between reliability and validity in optimizing our assessment.

This tension posed a real dilemma in planning our assessment scheme. We want assessment to be valid; that is, we want assessment to test that the outcomes have been achieved. We also want assessment to be reliable; that is, we want agreement between assessors. One way to boost reliability is to drive down the number of separate elements that must be passed. However, a valid scheme for professional assessment and accreditation will often require several distinct abilities to be demonstrated. For example, how would students feel knowing that their tutor was competent at planning lessons, at teaching, at record keeping and at reflecting on his or her practice, but not at assessment?

In trying to reconcile these conflicting pressures for reliability and for validity, we looked again at the set of outcomes we assess. We agreed that planning teaching sessions, teaching, assessing student work, and monitoring and evaluating teaching are vital and distinct outcomes. Keeping records of
teaching, reflecting on practice and continuing professional development are also important, but perhaps they could be subsumed by or combined with the others. And coping with the demands of the job is tricky to assess convincingly, and so might be dropped. Thus we might consider reducing to four the number of separate outcomes that must be passed, subject of course to external accreditation requirements. This, of itself, should improve reliability.

The requirement that each outcome be underpinned by specified values has added complexity to the assessment. But the idea that teaching is a values-based as well as a skills-based activity is one we hold dear, and anyway is reflected in the national accreditation frameworks to which we are committed. We thought about how we might simplify assessment, and hence reduce the scope for inter-assessor disagreement, without losing the underpinning values. One suggestion was that we might require that candidates show how, say, at least two values of their choice from the list underpin their attainment of each outcome, and also how, say, each value underpins at least two of the outcomes. This would reduce the number of assessment judgements to be made while maintaining the importance given to both the outcomes and the values. Or we might ask for separate accounts, first of how each outcome is demonstrably achieved, and then of how each value demonstrably underpins practice.

Finally, we analysed the assessment data in more detail. We asked: ‘Where is there the most disagreement between the two assessors?’ We looked at disagreement on the assessment of the outcomes and of the underpinning professional values. We found a particularly high level of disagreement on the outcome ‘reflection on practice’. Why? One reason was that this outcome is actually three separate outcomes. Course participants must show that they have 1) reflected on their work, 2) as a result, analysed the areas in which they need to develop further, and 3) on this basis planned their further professional development. In looking at individual portfolios and at assessors’ comments, we saw that assessors were disagreeing in particular about how explicit the needs analysis was, and about what exactly constituted a development plan. We clarified advice to candidates and to assessors, and we hope for improvement in pass rate and in assessor agreement on this outcome.

We also found a high level of disagreement about the underpinning professional value ‘commitment to equal opportunities’. Not surprising, perhaps – it is a highly contested concept. However, we had hoped for better. We have now clarified advice to course participants and tutors on this. We have also decided to give to course participants materials on equality of opportunity that were prepared for a later course in the series.

It is too early to tell whether our interventions will have the desired effects. We are hopeful that they will. At the same time, one heretical question has occurred to us. Need we be excessively worried if we can’t increase inter-assessor reliability, and thus that more people will fail the course than we might wish? Is failure on this course all that serious?
Of course failure is disappointing, but what actually follows from a failure? On this course, what follows is that the course participant who is judged to have failed is invited to do further work on the outcome(s) he or she has failed to demonstrate, and to resubmit. Such candidates receive detailed written feedback on the ways in which their work on the failed outcomes was judged unsatisfactory. Resubmission rates, and success rates at resubmission, are high. So maybe all's well that ends well, for the teachers and later for their well-taught students!

What do you think of the conclusions that were reached and the ideas for increasing inter-assessor reliability? What are the lessons from this case for your own assessment practice?

CASE REPORTERS' DISCUSSION

As already mentioned, our experience highlighted for us the tension between validity and reliability in assessment. Both are crucial elements, but we chose in this instance to focus on reliability. We claim that we could improve reliability by decreasing the number of separate elements to be judged by the assessors, where all have to be passed for completion of the course as a whole. A brief look at the relevant arithmetic, using our own course as an example, shows why this would be true.

Performance in our course is assessed on each of seven outcomes. Overall, we found that the level of agreement between the two assessors on whether or not each outcome had been passed was 87 per cent. Given that each and every outcome has to be passed for the course to be passed, how much agreement would we expect on whether a particular candidate had passed or not? The answer turns on how independent a lecturer’s performance on each outcome might be from his or her performance on every other outcome.

There is a spectrum of possibilities here. At one end of this spectrum, teaching is seen as a single unitary ability. This view would lead us to expect very high correlations among a candidate’s performances on all outcomes, and in turn would lead us to predict an overall pass/fail agreement between the two assessors of, again, something like 87 per cent. At the other end of the spectrum, perhaps each of the seven course outcomes describes a capability wholly independent of each of the others. If inter-assessor agreement is assumed to be 87 per cent for any outcome, then the chance that the assessors will agree on all seven outcomes, assuming that the agreements are randomly dispersed, is (0.87)^7 = 38 per cent. The approximately 60 per cent agreement that we found is between these extremes, suggesting that various skills of a teacher are somewhat, but far from wholly, interdependent.

This reasoning does not just apply to assessing portfolios. It applies to any situation with more than one element that must be assessed. The greater the
number of separate parts of an assessment that must be passed for a course to be passed, the more are magnified the effects of any unreliability of assessment on the separate elements being assessed. Again taking our own course as the example, what might happen if, as we suggested, we reduced the number of outcomes that must be passed from seven to four? Assuming the same inter-assessor pass/fail agreement on each outcome (0.87), and again assuming performance on these outcomes to be independent, we might anticipate an overall pass/fail agreement of \((0.87)^4\) (- 57 per cent). The chances of disagreement are thus lower with four outcomes than with seven. Furthermore, if there were a similar degree of alignment of pass/pass and fail/fail decisions as we had previously observed, we would predict a significantly higher percentage of agreement on course pass/fail than the 60 per cent we achieved with the seven outcomes.

As shown in the case, the added complexity of including values as well as outcomes might be approached in the same manner, i.e. reducing to a minimum the number of discrete items to be judged, all of which have to be passed. However, a further step can be the sort of analysis we undertook of those individual areas in which there was the greatest degree of disagreement between assessors. This led us to identify outcomes that actually comprised more components than we had thought, and values in which we should provide better information and guidance to candidates and/or assessors. More generally, this kind of detailed discrepancy analysis enables prioritization, on a rational basis, of the effort to be applied to improving assessment on courses. A similar identification of the outcomes or questions on which students score less well might tell where effort might usefully be applied to the materials or the teaching of a course.

Again, it is too early to know if the particular changes we made are going to be successful or not. However, the notion of continually monitoring, evaluating, reflecting on and attempting to improve both teaching and assessment is both something that is encouraged in our course materials and something we believe in and model ourselves. We do practice what we teach!

Acknowledgement

The production of an Open University course involves a large team of people. I (DB) am very grateful to colleagues in, particularly, the Centre for Higher Education Practice and The Open University's Institute of Educational Technology for all their work in producing the course analysed here, and to the tutors and course participants for their commitment and effort.

A very thorough analysis of the reliability of assessment of 53 portfolios on an Open University course on teaching in higher education, together with detailed implications for the improvement of assessment and a theoretical conceptualisation of assessment as the search for agreement among the various parties to the assessment.
The Reliability of Assessment by Portfolio on a Course to Develop and Accredit Teachers in Higher Education

DAVID BAUME
Centre for Higher Education Practice, The Open University, UK

MANTZ YORKE
Centre for Higher Education Development, Liverpool John Moores University, UK

ABSTRACT Portfolios are widely used to document and assess professional development. They are used to assess university teachers on courses run by the UK Open University. These portfolios are assessed twice, by trained assessors, against a detailed set of requirements which include learning outcomes and underpinning values. A detailed analysis was undertaken of the assessment judgements involved in the assessment of 53 such portfolios. Inter-rater reliability data are reported. These results are compared with those of other studies on portfolio assessment. Consideration is given to appropriate measures of the reliability of assessment, and to some effects of the structure of assessment and of the rules for combining scores on the reliability of assessment. Some implications for practice are explored.

Introduction
Portfolios and Their Use in Higher Education Teacher Development

Portfolios are becoming increasingly important in professional development; for example, Seldin (1997) indicates their significance for the discipline of education. The 'authenticity' of portfolios (being derived from experience in real-life settings), and hence portfolios' potential for validity, are seen as having advantages over exercises and examinations that are related less closely to the demands and realities of professional life.

There is variation in expectation as to what should be included in a portfolio (Stecher, 1998; Simon & Forrette-Giroux, 2000). However, a portfolio typically includes evidence drawn from practice. Crucially, it usually also contains reflective commentary, or (in the terminology used in the course considered here) a claim, in which the course participants show how they have interrogated their experience and related their practice and understandings to cognate evidence from the literature and elsewhere. It is typically expected that the portfolio will be scholarly, and that insights will go beyond a quotidian pragmatism to connect with relevant theoretical constructs. The assumption here is that theory is an important component for the bridges being built between practice in different contexts (Taylor, et al., 1999).

ISSN 0307-5079 print; ISSN 1470-174X online/02/010007-19 © 2002 Society for Research into Higher Education
DOI: 10.1080/030750701200999340
Such portfolios are used widely in programmes for the education and accreditation of teachers in higher education. The UK Staff and Educational Development Association (SEDA) has since 1992 recognised some 60 such programmes in the UK, Australia, New Zealand, Hong Kong, Singapore and Sri Lanka, and has accredited some 1700 teachers who have successfully completed recognised programmes (Baume & Baume, 1996; SEDA, 2000). Two linked qualifications have been established. Accredited teacher status is gained by those who design and run courses and programmes of studies, typically but not necessarily full-time university lecturers. The associate teacher qualification is designed for those whose work involves mainly teaching, typically part-time teachers.

SEDA's approach to accreditation adds two further elements to the evidence and the reflective commentary or claim already discussed. The first of these elements is a statement of the outcomes that a course participant must achieve in order to be accredited. These outcomes are relatively open accounts of what a teacher does—they include planning courses and classes, teaching and assessing student work. The second element is an account of the principles and values that must be shown to underpin the work of a teacher. These include an understanding of how students learn, commitments to student learning and to scholarship, and a concern for equality of opportunity. The outcomes and the underpinning principles and values for the course considered here are shown in Table I.

The UK National Committee of Inquiry into Higher Education (the Dearing Committee) recommended that 'over the medium term, it should become the normal requirement that all new full-time academic staff with teaching responsibilities are required to achieve at least associate membership of the Institute for Learning and Teaching in Higher Education, for the successful completion of probation' (NCIHE, 1997, Recommendation 48). The Institute for Learning and Teaching (ILT) was set up in the summer of 1999 to implement this and other Dearing recommendations. The ILT has added a further dimension to accreditation—required underpinning knowledge (ILT, 2000). During its first year, the ILT accredited SEDA-recognised courses and allowed SEDA-accredited teachers into membership, acknowledging the considerable commonality between ILT and SEDA approaches to accreditation.

Despite the wide and growing use of portfolios for professional accreditation, the reliability of portfolio assessments has been little addressed. There has been a steady trickle of papers that cast doubt on the reliability of essay-type assessments (see Brown et al., 1997; Gibbs & Rowntree, 1999). We anticipated that some of the strictures relating to the assessment of essays would transfer to the realm of portfolios.

The Course and Its Assessment

The course whose assessment is considered here—'Teaching in Higher Education' (H851)—was developed by the Centre for Higher Education Practice (CeHEP) at the Open University (OU) in the UK. It is designed to develop and accredit the teaching abilities of those who teach in higher education (those taking the course are in this paper called 'course participants' or simply 'participants', to distinguish the lecturers who take the course from the students whom they teach).

The course attracts 30 credits at master's level (180 credits are required for a UK master's degree). It is accredited by SEDA as leading to SEDA Associated Teacher status. A further course, 'Course Design in Higher Education', of similar size, when combined with 'Teaching in Higher Education', leads to a Postgraduate Certificate in Teaching and Learning in Higher Education, to SEDA Accredited status and to eligibility for membership of the Institute for Learning and Teaching.
TABLE I. Learning outcomes and underpinnings for H851, ‘Teaching in Higher Education’ slightly adapted from the SEDA associate teacher qualification [SEDA, 2000]

<table>
<thead>
<tr>
<th>Outcomes</th>
<th>Underpinning values</th>
</tr>
</thead>
<tbody>
<tr>
<td>You should show that you can:</td>
<td>You should show how these principles and values underpin your attainment of the outcomes:</td>
</tr>
<tr>
<td>1. Plan teaching sessions</td>
<td>1. An understanding of how students learn</td>
</tr>
<tr>
<td>Design teaching sessions from a course outline, document or syllabus.</td>
<td>All teaching and academic administration should be informed by an understanding of how students learn and the conditions and processes that support student learning. Underpinning outcomes 1, 2, 3, 4 and 7</td>
</tr>
<tr>
<td>This involves choosing teaching methods appropriate to:</td>
<td>2. A concern for students’ development</td>
</tr>
<tr>
<td>the group of learners:</td>
<td>Helping students to learn must begin with a recognition that all students have their own individual learning needs and bring their own knowledge and resources to the learning process. Work with students should empower them and enable them to develop greater capability and competence in their personal and professional lives. Underpinning outcomes 1, 2, 3, 4 and 7</td>
</tr>
<tr>
<td>the mode of study;</td>
<td>3. A commitment to scholarship</td>
</tr>
<tr>
<td>the subject material;</td>
<td>At the base of professional teaching is an awareness and acknowledgement of the ideas and theories of others. All teaching should be underpinned by a searching out of new knowledge—both about the subject/discipline and about good teaching and learning practice. All teaching should also lead to students developing a questioning and analytical approach. Underpinning outcomes 4 and 7</td>
</tr>
<tr>
<td>the resources available;</td>
<td>4. A commitment to work with and learn from colleagues</td>
</tr>
<tr>
<td>and the learning outcomes</td>
<td>Much of an academic’s work is carried out as part of a team made up of teaching staff and academic support staff. The colleagueship and support of peers is as important as individual academic excellence. Underpinning outcomes 4, 5 and 7</td>
</tr>
<tr>
<td>2. Teach</td>
<td>5. The practising of equal opportunities</td>
</tr>
<tr>
<td>Use two appropriate teaching and learning methods from:</td>
<td>Teachers must be concerned that students have equal opportunities, irrespective of disabilities, religion, sexual orientation, race or gender. So, everything that teachers do should be informed by equal opportunities legislation, by institutional policy and by a knowledge of best practice. Underpinning outcomes 1, 2, 3, 4 and 7</td>
</tr>
<tr>
<td>making presentations (e.g. lectures, demonstrations);</td>
<td>6. Continuing reflection on professional practice</td>
</tr>
<tr>
<td>facilitating group learning (e.g. through seminars, discussion groups, projects); working with individual learners; facilitating practicals or laboratory classes. You must also include evidence of using appropriate learning technologies.</td>
<td>Teachers should reflect on their intentions and actions and on the effects of their actions. They try to understand the reasons for what they see and for the effects of their actions. They thus continue to develop their understanding and practice and therefore inform their own learning. Underpinning outcomes 2-7 inclusive</td>
</tr>
<tr>
<td>3. Assess student work</td>
<td></td>
</tr>
<tr>
<td>Mark or grade, and give feedback on, student work.</td>
<td></td>
</tr>
<tr>
<td>4. Monitor and evaluate teaching</td>
<td></td>
</tr>
<tr>
<td>Monitor and evaluate your own teaching, using self, peer and student feedback.</td>
<td></td>
</tr>
<tr>
<td>5. Keep records</td>
<td></td>
</tr>
<tr>
<td>Keep appropriate records of your teaching support and academic administrative work.</td>
<td></td>
</tr>
<tr>
<td>6. Cope</td>
<td></td>
</tr>
<tr>
<td>Develop personal and professional coping strategies appropriate to the constraints and opportunities of your institutional setting, to manage adequately your time and operate successfully within available resources.</td>
<td></td>
</tr>
<tr>
<td>7. Continue your professional development</td>
<td></td>
</tr>
<tr>
<td>Reflect on your own personal and professional practice and development, assess your future development needs, and make a plan for your continuing professional development.</td>
<td></td>
</tr>
</tbody>
</table>
Course content materials deal with the practice and theory of teaching and learning in higher education. Course process materials guide participants in the development of their teaching abilities and in building the portfolio on which they are assessed. Guidance is also given to course participants on the assessment standards and process of the course. This guidance is given both through precept (detailed written accounts of criteria and standards) and examples (drafts of course participants' claims that an outcome has been met, with tutor commentary on the claims). Participants are tutored online using the FirstClass conferencing system, through which they also undertake online activities. Participants receive feedback from their tutor on draft sections of their portfolio.

Assessment on the course is a fine-grained process. To pass the course, participants need to show that they have achieved each of the seven course outcomes. Further, some of these outcomes have more than one component. For example, Outcome 7 requires the teacher to reflect on their practice, analyse their future development needs as a teacher and plan their further development. Further detail is added to assessment by the requirement that the underpinning principles and values demonstrably inform achievement of the outcomes, and by the specification of word and page limits.

Hundreds or thousands of people may take an Open University course. This in turn requires large numbers of tutors and assessors. A systematic approach is needed to making the course assessment process as even-handed and as manageable as possible. To this end, the course team for H851 has specified the assessment requirements, including which outcomes which must be shown to be underpinned by which values, in great detail.

Assessors and course team additionally undertake a process before each round of assessments which is designed to surface and then reduce differences in assessment judgements. Each assessor receives and assesses the same one or two portfolios. They then attend a coordination meeting at which they share assessment judgements and discuss and agree steps to reduce these disagreements in the assessments they are about to conduct. Notes of the meeting and the decisions are circulated to assessors.

A total of 75 judgements is recorded for each portfolio, including the overall pass/fail judgement. Different scales are used for the different kinds of elements assessed (see Table II). Combination rules are applied to achieve a judgement, first for each outcome, based on its elements, and then on the course as a whole, based on the seven outcomes.

- An outcome is judged to have been passed when each element has been well or just achieved, with the exception that a single 'not quite achieved' in one element in an outcome may still lead to a pass on that outcome.
- For the course to have been completed successfully, all seven outcomes must be achieved—at least at the level of a bare pass.
- If there is agreement between the two assessors regarding an overall pass or fail, then this judgement stands. If there is disagreement, a third marker is used, usually a course team member.

A course participant who is judged to have failed one, two or three outcomes is invited to undertake further work on them, and is given guidance on what elements in the outcomes failed need further work before they resubmit the portfolio. More than three outcomes failed means that the whole course must be retaken.

Detailed guidance is given to course participants and assessors on making these judgements, with examples of course participants' work and assessors' comments. The rationale for the assessment structure is discussed later in this article, as are the implications of this assessment process for course participants' results.
<table>
<thead>
<tr>
<th>Categ.</th>
<th>No.</th>
<th>Contents</th>
<th>How judged?</th>
<th>Grading scale</th>
<th>Rules</th>
</tr>
</thead>
<tbody>
<tr>
<td>A &amp; B</td>
<td>21</td>
<td>Technicalities, relating to, for example, portfolio word and page count, the presence of items including cv and authentication of teaching and portfolio.</td>
<td>By count or check. Each judgement is independent</td>
<td>Two points: Yes and No</td>
<td>Exceeding the word or page limit does not cause a fail. However, assessors are asked to ignore material beyond the limit.</td>
</tr>
<tr>
<td>Ci</td>
<td>20</td>
<td>Elements of the portfolio, usually subdivisions of the outcome.</td>
<td>By academic judgement. Each judgement is independent</td>
<td>Four points: Well achieved, just achieved, Not quite achieved, Not achieved</td>
<td></td>
</tr>
<tr>
<td>Cii</td>
<td>26</td>
<td>Assessments of the six underpinning values, from two to six specified for each outcome.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D</td>
<td>7</td>
<td>Course outcomes</td>
<td>By combining judgements on elements (as at right)</td>
<td>Five points: Outstanding, Clear or Bare pass, Bare fail or (Clear) Fail</td>
<td></td>
</tr>
<tr>
<td>E</td>
<td>1</td>
<td>Overall result (pass/fail)</td>
<td>By combining judgements on outcomes (as at right)</td>
<td>Two points: Pass, Fail</td>
<td>If any item in an outcome is not achieved, the outcome is failed. If just one item in an outcome is not quite achieved, the outcome should be judged to have been achieved. If any outcome is judged a bare or clear fail, the course is failed. If 1 to 3 outcomes are failed, resubmission is possible. If more than 3, the course must be retaken.</td>
</tr>
<tr>
<td>Total</td>
<td>75</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Any summative assessment process needs to be both valid and reliable. We make the presumption here that the assessment of this course is likely to be generally valid, in that the assessment process assesses explicitly whether course participants have achieved the stated learning outcomes of the course, and have done so in a way that is demonstrably informed by the underpinning values. However, the validity of the espoused assessment approach would be weakened if the assessment-in-use gave rise to problems. Divergences between assessors, i.e. indications of unreliability, would therefore cause us to look again at validity.

The primary research questions for us, then, were 'What is the reliability of portfolio assessments?', and subsequently, 'Why are the reliability figures as they are?' We had available to us detailed assessment data, and we have used these data to produce a substantial case study. However, we believe that the issues raised by our analyses are relevant to a much wider community of assessors on academic and professional courses.

The OU Examinations and Assessment Office provided us with detailed results for 53 assessments of portfolios from two cohorts of participants on the course. Data were available for both first and second markers, making it possible to compute the inter-assessor reliability. Although the two sets of portfolios were assessed in different years, we felt that we were justified in bringing the original pairs of assessments together for the purposes of analysis, since the procedure was common and there had been no changes of substance during the period at issue.

Results and Discussion

Success Rates

The profile of agreed success rates in respect of the elements, underpinning values and course outcomes is shown in Table III. Table III shows that, by and large, the participants on H851 are successful in attaining the range of outcomes specified. Where graded judgements are concerned, performances are concentrated at the 'well achieved' level for elements and underpinning values, and at the 'clear pass' level for course outcomes. In part, the high levels of performance are attained because participants receive feedback from their tutors on draft sections of their portfolios before submitting the final version for assessment. However, despite the efforts of the course team to be explicit with assessors and course participants about assessment requirements, course team and assessors retain some tacit 'guild knowledge' (Sadler, 1989). Tutor comment helps the course participants to appreciate this tacit knowledge, and hence to elaborate their understandings of what the course expects.

Technical requirements gave rise to very few failures or significant disagreements. The tasks that require qualitative judgements are of greater interest here. Table III shows that, whilst the success rate on individual outcomes (in pass/fail terms, which is the primary consideration for a course participant) remained above 90%, the success rate dropped to 61% as assessments were brought together for the overall assessment. The combination rules shown in Table II have this consequence, and are considered further below. At this stage, we should note that a single clear fail on an individual component therefore carries through to produce an overall fail for the portfolio.

Assessor Agreement

Our consideration of the outcomes of the assessments of the portfolio components into six categories is based on the categorisation shown in Table II. We defined three degrees of
### Table III. Percentages of passing grades on elements, underpinning values and course outcomes for 53 portfolios (rounding accounts for inconsistency between categories of pass and total pass)

<table>
<thead>
<tr>
<th>Percentages of assessment in each category</th>
<th>Well achieved</th>
<th>Just achieved</th>
<th>Not quite achieved</th>
<th>Not achieved</th>
<th>Total pass</th>
<th>Total fail</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Elements and values (Ci and Cii), n = 46</strong></td>
<td>66</td>
<td>27</td>
<td>6</td>
<td>1</td>
<td>93</td>
<td>7</td>
</tr>
<tr>
<td><strong>Course outcomes, n = 7</strong></td>
<td>17</td>
<td>48</td>
<td>28</td>
<td>3</td>
<td>92</td>
<td>8</td>
</tr>
<tr>
<td><strong>Overall result</strong></td>
<td>61</td>
<td>39</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
agreement. When two assessors award the same grade, we call this ‘exact agreement’. When they award either the same grade or adjacent grades, we call this ‘close agreement’. When they agree whether the item being assessed should pass or fail, we call this ‘pass–fail agreement’. Table IV shows (as applicable) the percentages of these three kinds of agreement—exact, close and pass–fail—for each of the six categories for the 53 portfolios.

We chose to focus on inter-assessor agreement in respect of categories Ci, Cii and D, since to do so allowed us to compare our findings with those from the literature on portfolio assessment, in which portfolios were typically assessed on 4- to 6-point scales. It will be recalled from Table II that Ci and Cii are assessed on a 4-point scale, and D is assessed on a 5-point scale (the close/exact distinction is irrelevant in respect of categories A, B and E, since these categories are assessed on a simple yes/no or pass/fail basis). The row ‘percentage pass/fail agreement’ only adds information for column D, course outcomes. The row is included because passing or failing any outcome is a determinant of the overall result on the course: shadings of passing or failing are irrelevant at this point.

The inter-assessor reliability was computed at the three degrees of agreement described earlier: the percentages of exact agreement, close agreement and pass/fail agreement. We believe that the percentage agreement is a more appropriate measure than the inter-assessor correlation (which is also found in the literature). The latter will produce a high value when the assessors’ ratings have similar patterns but at different levels (e.g. when, albeit on rare occasions, one marker is systematically ‘tougher’ than another, despite a training programme). The Pearson r statistic is probably inappropriately used in a number of studies, since the data are ordinal rather than interval in character: a non-parametric measure would be preferable (see discussion in, for example, Boyatzis, 1998).

How should we evaluate the extent of agreement observed? Assuming complete randomness of judgement, the theoretical distribution of exact and close agreements between two assessors can easily be calculated. For a 4-point scale, there is a 25% chance of full agreement and a 63% chance of a close agreement; for a 5-point scale, the respective figures are 20% and 52%. In general, for an N-point scale, the chance of exact agreement is 1/N, of close agreement \((3N - 2)/N^2\). The percentages of agreements observed should be interpreted with reference to these statistics (note that we are simply considering agreement, not correctness of judgement).

In order to test whether the distributions actually observed were significantly different from random, the theoretical distributions for the 4-point scale (Ci and Cii) and 5-point scale (D) were calculated, assuming randomness in grading. Kolmogorov-Smirnov one-sample tests were run in respect of each of the individual items making up Ci, Cii and D. The one-tailed test (Conover, 1971) was used since, a priori, the presumption existed that the level of agreement would be higher than that afforded by chance alone. The results showed that the distributions are, statistically, considerably distinct from random, and therefore that there is real agreement among assessors on the elements of assessment and on the assessment of outcomes.

The relatively high level of agreement at the level of course outcomes, shown in the percentages of close agreement contained in Table IV may be attributable in part to the assessor training and to the coordination process described earlier. Another contributing factor may be the detail with which curricular expectations are both expressed as principles and illustrated with examples, in the course materials and in assessor briefing. Sadler (1989) argues for the importance, for the recipient, of having both descriptive statements and exemplars of expected outcomes, and his point is supported empirically (Wolf, 1995). Polanyi (1962, pp. 53–54) made a stronger case for the use of examples: ‘Connoisseurship ... can be communicated only by example, not by precept’. It seems likely that the assessment
<table>
<thead>
<tr>
<th>Assessments per portfolio</th>
<th>A</th>
<th>B</th>
<th>Claim within size limit</th>
<th>Evidence within page limit</th>
<th>Elements</th>
<th>Underpinning values</th>
<th>Overall result</th>
<th>Cii</th>
<th>Underpinning values</th>
<th>Overall result</th>
<th>D</th>
<th>Course outcomes</th>
<th>Overall result</th>
<th>E</th>
<th>Overall result</th>
</tr>
</thead>
<tbody>
<tr>
<td>Assessments n = 53 portfolios</td>
<td>97</td>
<td>96</td>
<td>64</td>
<td>55</td>
<td>86</td>
<td>85</td>
<td>87</td>
<td>60</td>
<td>86</td>
<td>86</td>
<td>86</td>
<td>60</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Percentage exact agreement</td>
<td>97</td>
<td>96</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Percentage close agreement</td>
<td>64</td>
<td>55</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Percentage pass/fail agreement</td>
<td>86</td>
<td>87</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Reliability of Assessment by Portfolio
of the H851 course is more solidly underpinned than have been the assessments of portfolios in some other contexts.

There was a very high level of agreement between pairs of assessors regarding the technicalities. Since the criteria for technicalities are precisely drawn (e.g. maximum wordage for a claim), it is perhaps disappointing that there should be any disagreement at all. One source of disagreement is that some assessors appear to have ignored small infractions. As with success rates, the more interesting reliability data relate to the curricular components that require graded judgements.

Table IV shows that inter-assessor exact agreement is moderate for elements and values, and poor for course outcomes. However, when the data are reduced to a simple pass/fail division, the levels of agreement are much better. In other words, whereas assessors might disagree about the quality of a pass (or, occasionally, of a failing performance), they seem to have less of a problem with the 'cut' between pass and fail. As with success rate, when exact agreement data are worked through to the course outcome level, the percentage of exact agreements drops considerably. Table IV strongly hints that, with three grades of pass at course outcome level, the scope for relatively unimportant disagreements about the standard of a pass is quite high. It is not surprising, therefore, to find the percentage of exact agreement rising at the overall assessment level, since that level is only concerned with whether or not the course outcomes have been achieved.

The agreement rate of 60% regarding the whole portfolio is lower than any assessment regimen would like. We found examples where one of a pair of markers had marked much lower than the other. In the most extreme instance, one marker had marked 42 out of 74 components lower than the other. Perhaps not surprisingly, these markers disagreed about the overall result. There were also two cases in which the pattern of disagreement was reversed for different sets of elements and values, suggesting, perhaps, that assessors were differentially secure regarding their areas of expertise or simply that they had different calibrations for different outcomes.

We noted earlier that, where there is a significant disagreement between the two assessors for a portfolio (that is, a disagreement whether a course participant had passed or failed the course), a third assessor assesses the portfolio. The third assessor concentrates on the areas of disagreement between the first two assessors, and determines the grade to be awarded. Cresswell (1986) argues that assessment by a further assessor, who uses the same criteria and knows the originally-awarded mark(s), is the method most likely to reduce marking errors. Cresswell considers the case of a second rather than a third assessor marking, but much the same argument applies to the assessments discussed here.

The use of a third marker is, however, an extra resource burden that, for obvious practical reasons, needs to be minimised. Defining the expectations of the course more precisely and enhancing the training for assessors should help to reduce discrepancies between assessors, and hence the overall cost of assessment.

**Particular Discrepancies between Assessors**

Assessors are asked to judge whether candidates have shown how specified course outcomes are underpinned by specified course values, as indicated in Table I. Table V shows the number of occasions on which the two assessors disagreed (by more than one grade) on these judgements (cells are blank where the particular value is not required to be shown underpinning the particular value). These discrepancies in judgement between assessors indicate areas of difficulty in interpretation of the course's requirements (on the part of assessors and/or
course participants), and thereby indicate where particular attention is needed from the course team.

On the principle that discrepancy scores above the overall mean should be the primary focal points for attention, this analysis suggests that there are problems with the assessment of Outcome 7, and possibly also Outcome 5. Outcome 7 requires participants to review their teaching, analyse their teaching development needs and make a plan for their continuing professional development in order to address the needs identified. Outcome 5 requires the keeping of records of teaching support and academic administrative work.

The analysis also shows problems with two required underpinning values: Value 4, concern for equality of opportunity, and Value 6, reflection. These discrepancy scores have already been used to direct work by the course team, aimed at enhancing the briefing given to course participants and assessors regarding these topics, with a view to increasing the reliability with which they are assessed. This is considered further in the section on 'using the data to improve assessment practice'.

How Do the Data Compare with Those from Other Studies?

Reliability studies of portfolios of the performances of schoolchildren (Herman et al., 1993; Koretz et al., 1993; LeMahieu et al., 1995; Wolfe, 1996; Supovitz et al., 1997; Heller et al., 1998), college students (Nystrand et al., 1993) and medical general practice trainers (Pitts et al., 1999) indicate respectable levels of inter-rater agreement (Table VI). These inter-rater agreements have been typically achieved in circumstances in which there was a template of defined outcomes or criteria against which to judge. Judgements become difficult to make when there is insufficient information: for example, 'outsider' judges of children's performances found greater difficulty than did the children's teachers in making judgements, presumably because the teachers had a fuller knowledge of their pupils and could hence interpolate the missing data (Supovitz et al., 1997). Further, bias is possible when the assessor picks up from the portfolio cues about the assessee (Howell et al., 1993).

The studies of reliability suggest that, whilst it may be possible to secure a reasonable level of inter-rater agreement in the assessment of portfolios, the underlying 'scatter' of gradings (evidenced in the Pearson r data) could be tightened up. It can also be concluded from these studies that reliability is enhanced when there are explicit outcome standards against which to judge, and when there are clear and unambiguous performance data upon which to exercise that judgement.

Nystrand et al.'s (1993) work indicates that reliability may be higher when assessors grade a set of portfolios by taking one element at a time, and grading all students on that element before moving to the next element, than when they work their way through one portfolio before turning to the next. The 'element by element' approach may not be available for the OU courses or the ILT accreditation process, and hence all that is likely to be possible is to maximise the reliability of assessments made seriatim through individual portfolios.

We compare in Table VII the data from the current study with those from the work cited earlier. More data are needed to draw any strong conclusions. However this table at minimum suggests that (a) reliable portfolio assessment is not easy to achieve and (b) the reliability of the assessment on the CeHEP course described here stands comparison with evidence from other studies—at least, at the level of the assessment of individual elements and outcomes. A complicating factor in making comparison of reliability is the three-layer assessment scheme adopted in the CeHEP course, with rules governing the combination of results, first from elements to outcome and then from outcomes to overall judgement. This
<table>
<thead>
<tr>
<th>Underpinning value</th>
<th>Course outcome</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. How students learn</td>
<td>5</td>
<td>4</td>
<td>5</td>
<td>5</td>
<td></td>
<td>12</td>
<td>6.2</td>
</tr>
<tr>
<td>2. Concern for student development</td>
<td>7</td>
<td>7</td>
<td>5</td>
<td>5</td>
<td></td>
<td>14</td>
<td>7.6</td>
</tr>
<tr>
<td>3. Scholarship</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td>8</td>
<td></td>
<td>6.5</td>
</tr>
<tr>
<td>4. Colleagueship</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>6</td>
<td>10</td>
</tr>
<tr>
<td>5. Equal opportunities</td>
<td>9</td>
<td>9</td>
<td>11</td>
<td>15</td>
<td></td>
<td>12</td>
<td>11.2</td>
</tr>
<tr>
<td>6. Reflection</td>
<td>12</td>
<td>9</td>
<td>6</td>
<td>10</td>
<td>8</td>
<td>8</td>
<td>8.8</td>
</tr>
<tr>
<td>Mean</td>
<td>7.0</td>
<td>8.0</td>
<td>7.5</td>
<td>6.8</td>
<td>10.0</td>
<td>7.0</td>
<td>10.7</td>
</tr>
</tbody>
</table>

Note: discrepancies are instances of disagreement of more than one grade between the two assessors. The discrepancy rates for judgements of values underpinning outcomes from the assessment of 53 portfolios are presented here. There are 16 blank cells because not all outcomes are required to be underpinned by all values (see also Table 1).
TABLE VI. Reliability data from other studies on portfolio assessment

<table>
<thead>
<tr>
<th>Study</th>
<th>Reliability measure</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Heller et al. (1998)</td>
<td>Percentage of exact agreement ranged from 48 to 63, of agreement within one point from 91 to 100. Inter-rater reliability coefficients ranged from 0.53 to 0.83.</td>
<td>ns ranged from 5 to 13, total ( n = 84 ). Involved holistic ratings of portfolios from Grade 4 and Grade 8 pupils.</td>
</tr>
<tr>
<td>Herman et al. (1993)</td>
<td>Inter-rater agreement ranged from 89% to 100%, between pairs from 3 raters. Pearson r values ranged from 0.41 to 0.94.</td>
<td>Ratings cover whole portfolios and also components. One grade difference taken as agreement.</td>
</tr>
<tr>
<td>LeMahieu et al. (1995)</td>
<td>Inter-rater correlations ranged between 0.74 and 0.87*.</td>
<td>Middle school and secondary level writing portfolios.</td>
</tr>
<tr>
<td>Koretz et al. (1993)</td>
<td>Spearman rho correlations between raters around 0.60 for overall portfolio ratings.</td>
<td>Ratings for components of portfolios were lower.</td>
</tr>
<tr>
<td>Nystrand et al. (1993)</td>
<td>Inter-rater agreement (ns vary from 7 to 109) on portfolio elements ranged from 19% to 71%. Pearson r values ranged from — 0.35 to 0.66.</td>
<td>Portfolios assessed sequentially; 1 grade difference taken as agreement.</td>
</tr>
<tr>
<td>Supovitz et al. (1997)</td>
<td>Inter-rater agreement (ns vary from 48 to 493) on portfolio elements ranged from 53% to 79%. Pearson r values ranged from — 0.44 to 0.86.</td>
<td>Items in portfolios assessed across all assesses' elements rather than portfolios assessed as wholes. Same criterion of agreement.</td>
</tr>
<tr>
<td>Wolfe (1996)</td>
<td>Inter-rater correlations — 0.04 to 0.55, 0.47 to 0.79, and 0.46 to 0.96 for science, language arts and mathematics work samples respectively. Respective exact agreements (percentages) were 33—64, 34—61 and 43—91; agreements no worse than one point different were 87—98, 80—93 and 80–100.</td>
<td>Secondary school pupils' work.</td>
</tr>
<tr>
<td>Pitts et al. (1999)</td>
<td>Pass–fail agreement on assessment by 8 assessors of portfolios by 12 candidates on seven attributes ranged from 68% to 89%. Pass–fail agreement on a single overall assessment judgement was 75%. Reliabilities were reported using the kappa statistic.</td>
<td>Portfolio prepared by UK medical general practice trainers.</td>
</tr>
</tbody>
</table>

* Koretz (1998) argues that these figures are inflated due to inappropriate use of the Spearman–Brown prophecy formula in calculating reliability. According to Koretz, the real figures are 0.60–0.67 and 0.71–0.77 for high school and middle school portfolios respectively.

inevitably magnifies any unreliability at the level of the elements. This issue is considered further below.

Another approach can be used to compare the current study with Pitt's data, using the kappa-like statistic (k) (Davies & Fleiss, 1982), as shown in Table VIII. Accepting the normal guideline that \( k > 0.7 \), \( 0.7 > k > 0.4 \) and \( 0.4 > k \) respectively denote excellent, good and marginal agreement over the level of chance, then we see excellent close agreement at the
TABLE VII. Inter-assessor agreement from four studies of portfolio assessment

<table>
<thead>
<tr>
<th>Researchers</th>
<th>Exact agreement (%)</th>
<th>Close agreement (ratings no more than one grade apart) (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Heller et al. (1998)</td>
<td>48-63</td>
<td>91-100</td>
</tr>
<tr>
<td>Herman et al. (1993)</td>
<td>Not reported</td>
<td>89-100</td>
</tr>
<tr>
<td>Nystrand et al. (1993)</td>
<td>Not reported</td>
<td>19-79</td>
</tr>
<tr>
<td>Wolfe (1996)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| Science                      | 33-64               |
| Language arts                | 34-61               |
| Mathematics                  | 43-91               |
| Language arts                |                     |
| Mathematics                  |                     |

outcome level for the current study. Exact agreement on the rule-based outcome and overall judgements are poor. The current study also shows good exact agreement on elements and (just) on values. Pitts's study shows good agreement on the single overall assessment judgement (assessments on the course studied here are made on a multipoint scale; however, the data for agreement (whether exact or close) are dichotomous, being either 'agreement' or 'disagreement' within predetermined bands).

The Effects of Combination Rules

There are two polar views about the degree of interdependence between course outcomes. One of these holds that each course outcome involves a different ability. On this basis, we might anticipate that a participant's performance on an outcome would be substantially independent of their performance on each other outcome. The contrary view is that teaching is essentially a single activity, and involves a group of cognate abilities among which we might expect to see high correlation in the work of a teacher.

The former of these views is more strongly suggested than the latter by the observation that the overall pass rate is lower than the pass rate on individual outcomes. The overall agreed pass rate on outcomes is 92%; that on the course as a whole, 61% (were both performance and judgement on each of the seven outcomes to be quite independent, then in the simplest case the probability of an agreed pass would be \((0.92)^7 = 0.55\)).

A course participant must pass each outcome to pass the course. Two polar views can also be taken on the extent to which this outcome-by-outcome assessment is appropriate for judging the ability of a higher education teacher. One view is that a higher education teacher needs a range of particular abilities, here expressed as outcomes and underpinning values,

TABLE VIII. Inter-assessor agreement from two studies of portfolio assessment

<table>
<thead>
<tr>
<th>Researchers</th>
<th>Exact agreement (%)</th>
<th>kappa</th>
<th>Close agreement (%)</th>
<th>kappa</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pitts et al. (1999)</td>
<td>75</td>
<td>0.54</td>
<td>Overall (P/F)</td>
<td></td>
</tr>
<tr>
<td>(single judgements)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Current study</td>
<td>64</td>
<td>0.52</td>
<td>Elements (5 point scale)</td>
<td>86</td>
</tr>
<tr>
<td>(single judgements)</td>
<td>55</td>
<td>0.40</td>
<td>Values (5 point scale)</td>
<td>85</td>
</tr>
<tr>
<td>(rule-based combinations of judgements)</td>
<td>39</td>
<td>0.24</td>
<td>Outcomes</td>
<td>88</td>
</tr>
<tr>
<td></td>
<td>60</td>
<td>0.2</td>
<td>Overall (P/F)</td>
<td></td>
</tr>
</tbody>
</table>
Reliability of Assessment by Portfolio

and that each is a necessary component of a teacher's ability. The other view suggests that teaching is an integrated activity, and that a holistic judgement is more appropriate.

A principled accommodation between these poles can be offered. Some abilities (perhaps planning a course and lesson, teaching, marking and grading, and continuing to develop as a teacher) are clearly essential for the accreditation of practice. Others, though still important, could safely be worked towards and attained over time (provided that the participant recognised their importance and showed, for the purposes of initial accreditation, some movement towards attaining them).

Using the Data to Improve Assessment Practice

The data show that improvements in the assessment of this course are clearly possible. The analysis of discrepancies in assessment does not in itself resolve the underlying disagreements between the assessors. It does, however, suggest where different perspectives exist regarding what is to be valued in participants' portfolios, and hints, therefore, at problems with validity. It thus indicates priorities for the course team's further investigations and development efforts.

Closer inspection of the results for Outcome 7 showed better agreement on the first part, 'reviewing one's teaching', than on the second and third parts, 'analysing development needs' and 'drawing up a plan for continuing professional development to address these needs'. These are clearly two aspects of the course that need to be reviewed. The course team now gives greater emphasis in assessor briefings to identifying and judging needs analyses and plans for continuing professional development. Tutors are also asked to emphasise to course participants the importance of these elements.

The analysis also suggests particular problems with the assessment of two values. The meaning of 'concern for equality of opportunity' (Value 4) had already been a matter of sustained debate in the course team and among tutors. An engagement with equal opportunities is potentially wide ranging, since it is explicit in the course expectation that it will be addressed in respect of five of the seven course outcomes. Some course participants are faced inescapably from the outset of their teaching with the issue, because of the characteristics of those whom they are teaching or because of the emphasis placed in their institution on providing equality of opportunity. For others, the issue may be felt to be more theoretical than immediately practical. A lecturer in Earth Sciences may not be faced with learners who would have difficulty in accessing field study sites, to give a relatively unsubtle example. However, such a lecturer would be encouraged to reflect in their portfolio on how they would accommodate issues such as mobility, access, and ease of hearing outdoors, and also on why their course had not hitherto attracted students with disabilities. The assessor has to make a judgement of the course participant's performance with reference to the context of that performance. For the assessor, there will be some fuzziness in their understanding of both performance and context, complete understanding being unattainable. The course team, in debating equal opportunities, has considered whether successful efforts to avoid discriminatory practice (including language) met the equal opportunities requirement, or whether active efforts towards more inclusive practice were additionally required.

'Continued reflection on professional practice' (Value 6) has also proved a problematic value to assess in action, though rather less so than equality of opportunity. Another debate amongst the course team has explored whether reflection should simply look backwards, trying to make sense of what has happened, or whether the course requires reflection additionally to involve planning of how practice should change. A further debate asks to what extent reflection needs to be informed by theory.
Some of these problems have been addressed by course team decisions and consequent improved briefing of course participants and assessors, as already stated. The course team has also been able to draw on materials produced for a further course, on course design, which address some of the topics found problematic in assessment here. What this research has shown is that a fine-grained analysis of outcome data can pinpoint areas on which emphasis should be placed for the purposes of improvement.

Here is a final illustration of how close study of assessment can suggest where changes to course practice are necessary. Study of the claims provided by participants in this course under each outcome showed that they tend to use the statements of underpinning values as the basis of the organisation of their claims, interweaving commentary on their skills of teaching. This can mean that assessors have to ‘dig’ for the evidence relating to skill, whereas they have little difficulty in appreciating the teacher’s demonstration of the underpinning values. This need to dig may prejudice reliability in assessment. The course team is exploring alternative ways to structure the portfolio claims that will address this issue.

A Theoretical Consideration

Assessment is generally under-theorised, as an examination of a number of texts on assessment quickly reveals. It is not the purpose of this article to deal in any substantial way with this problem, but our study does throw a chink of light on to it.

Validity and reliability are typically related to outcomes, but the connection to outcomes diverts attention from the important contribution that the assessor makes. Reliable assessment requires shared knowledge (inter-subjectivity) on the part of the assessor and the assessee regarding the expectations laid down by the curriculum. There are thus relationships within which intersubjectivity is at risk—the relationships between an assessee, the assessors, and the curriculum. Even if the assessment approach is broadly valid, reliability is compromised to the extent that inter-subjectivity is under-developed.

We have noted, nevertheless, that in respect of some curricular components here, inter-subjectivity is likely to be more problematic than for the general run of courses: hence, there is a stimulus to initiate further developmental activity in respect of these. Let us explore this further.

Two main sets of actors are directly involved in assessment on any course, taking four main roles: the course participants, as assessees, and the lecturers who plan, teach and assess the course. The OU, like other institutions which run large courses, divides this role into two: the tutors who teach and assess the course, and the course team, who planned, wrote, and now run the course (which includes the moderation of assessment). The roles of others, such as the external examiner and institutional quality assurance processes, are not considered here. Assessment may be taken to measure, among other things, the extent of agreement among these sets of actors about what constitutes satisfactory performance on the course. (The justification for the assessments made is of importance here, since this makes transparent whether the actors share an understanding regarding what is at stake.) If understanding is to be maximised, then the following actions need to be mutually consistent.

- The course team defines and explains, for the benefit of both students, tutors and assessors, what it means by satisfactory work on the course.
- The tutors try to help course participants to understand and then deliver satisfactory work.
- Course participants try to produce what they consider to be satisfactory work.
- The assessors and the course team strive to assess whether or not work submitted is indeed satisfactory.
The work reported in this article can be seen as exploring the extent of agreement variously among course participants, tutors/assessors and course team. We discuss here two questions: 'What information do the three sets of actors share and not share?' and 'To what extent are these sets of actors members of communities of practice around assessment?' All three sets of actors involved in this course have ready access to a large set of information about assessment, both in higher education in general and on their course. This commonly available information directly related to assessment comprises: accounts of the course outcomes and underpinning values; practice guides and reader chapters on assessment; the course guide; the guide to preparing and presenting a portfolio; a guide to assessment on the course, written for participants and assessors; and simulated drafts of portfolios with tutor feedback. This is a substantial set of shared information.

Beyond having access to these materials, the three sets of actors also have shared experiences, chief among them the experience of teaching and assessing in higher education. Most tutors and some of the course team are accredited higher education teachers. Their shared experiences thus usually also include preparing for being assessed as a teacher, probably through a portfolio, to the same or a similar set of learning outcomes as those for the courses considered here. The actors will, collectively but differentially, share the explicit and tacit knowledge of professionals (see, in this respect, Eraut, 1994, chapter 7), and come close to constituting a 'community of practice' (Wenger, 1998). The level of shared experience is thus greater than that of a randomly-drawn group of higher education teachers, or of a typical group of staff and students in higher education.

Participants on the course produce draft sections of their portfolio, and receive feedback on these from their tutor. An aim of this feedback is to help course participants to achieve the necessary standard of teaching, as shown in the portfolio. This feedback from tutor to course participants is then monitored by a course team member. An aim of this monitoring is to help the tutor ensure that they are working to a common standard of judgement and of feedback. Both feedback processes, then, are aimed at supporting convergence about the standards of assessment required by the course.

There are thus plenty of similarities: shared access to much information, many shared if not identical experiences, many bases for agreement. What of the differences? First, current course participants have no access to the coordination process or to its decisions; in any event, such access would be too late to affect their (already submitted) work. However, the main decisions from coordination are fed into the assessment guidance for the next cohort. Second, and probably much more important, each course participant and every tutor has their own—different—experiences of teaching and assessment, on the course and in the rest of their working life, and makes their own sense and meaning of these experiences as well as of the course materials.

Conclusions

Our analysis of portfolios from a particular course (H851) has shown the following.

- The reliability of overall assessment on the course stands comparison with the spectrum of reported reliabilities for cognate assessments.
- The reliability of assessment at the outcome level is high, probably because of the extensive steps taken to achieve reliability.
- The cumulation of component assessments is in general very likely to reduce the reliability of overall judgements and increase the need for third marking, with associated costs.
- Detailed analyses of the kind reported here can lead to an enhancement of assessment
practice and, as a consequence, improvements in validity and reliability, and can further suggest priorities for this quality enhancement work.

Our work carries implications for the assessment of portfolios in a variety of educational settings. Of particular importance are the need for explicitness regarding the expectations placed on the assessee, and the need to ensure that all who are party to the assessment practice share an awareness of what is expected. If inter-subjectivity regarding an assessment is high, then the reliability of the assessment (at least) ought to be high.

One issue not resolved by this study is the proper relationship between holism and particularism in assessment. Another is the proper balance between components that must be passed and those regarding which a greater tolerance can be allowed. These matters must be tackled in the assessment of portfolios if the judgments made are to stand up to critical scrutiny.

Acknowledgements

We are grateful to the Open University in several respects. Student Services, via its Director, Professor David Sewart, funded Mantz Yorke at the University, during which this work was undertaken, and the experimental marking of portfolios. The Examinations and Assessment Office and its Director, Ben Palmer, supplied us with the assessment data in machine-readable form. The Course Team for H851, chaired by Kathy Nicoll, and the Programme Director, Dr Barbara Hodgson, both of the OU's Institute of Educational Technology (IET), supported and encouraged the work. Also in IET, Gordon Burt helped us to analyse the theoretical distribution of agreement between assessors. Within the Centre for Higher Education Practice, Research Fellow Martin Coffey gave generous support in data preparation and Eileen Goodyer and Sew Lan Hamilton gave valued secretarial support. Godfrey Pell gave statistical advice, and Carole Baume (the first presentation course team chair for H851) and Professor Graham Gibbs (Director of Research in the Centre) gave academic and moral support.

Correspondence: David Baume, Director (Teaching Development), Centre for Higher Education Practice, The Open University, Walton Hall, Milton Keynes MK7 6AA, UK; e-mail a.d.baume@open.ac.uk

REFERENCES


STAFF AND EDUCATIONAL DEVELOPMENT ASSOCIATION (2000) *The SEDA Teacher Accreditation Scheme* (Birmingham, SEDA); available at: http://www.seda.demon.co.uk/taccred.html


Publication 10  Baume, D. and Yorke, M. with Coffey, M. (2004). What is happening when we assess, and how can we use our understanding of this to improve assessment? *Assessment and Evaluation in Higher Education* 29(4): 451-477. ISSN 0260-2938

A very fine-grained experimental study of the reasons why assessors make the assessment judgements that they make and of the alternative decisions that they consider, again drawing implications for the improvement of assessment. The substantial differences in grades found to be awarded under real and experimental conditions are also explored.
What is happening when we assess, and how can we use our understanding of this to improve assessment?

David Baume1 & Mantz Yorke*2 with Martin Coffey3
1formerly of The Open University, UK; 2Liverpool John Moores University, UK; 3University of Leicester, UK

In an attempt to gain a fuller understanding of the basis of grading, ten assessors each assessed two portfolios drawn from the course archive which had been produced by participants on a course in teaching in higher education. Assessors gave a grade or judgement on each of a portfolio's 75 portfolio elements, reasons for each judgement they made, and recorded any alternative judgement they had considered. There were substantial differences between the judgements made during the study and those made when the portfolios were originally assessed. This article explores the differences between the original and experimental assessments, and then analyses reasons given by assessors for the judgements made. Beyond these particulars, the paper suggests some fundamental and problematic issues about the conduct and the reform of assessment, and offers a process which can be used to investigate and improve the quality, in particular the reliability, of any summative assessment.

Introduction

The complex and problematic nature of assessment has been addressed extensively (for example, see Brown et al., 1997; Heywood, 2000). However, the actual process of assessment has received rather less research attention. The investigation described in this paper is highly specific, as is the course whose assessment is studied. In particular, the assessment schema is very tightly defined, giving rise to the need for assessors to make large numbers of decisions on each portfolio. Baume and Yorke (2002) analysed archival data for this course, and found—inter alia—that certain components of the assessment demand seemed to be creating particular difficulties for assessors. The archival data, of course, gave no indication of how the assessors were attempting to deal with the challenges posed by the assessment process. This article presents findings from a follow-up study in which actual assessments are accompanied by assessors' commentaries on the process.

The original purpose of the investigation was to try to understand better what is
happening when a portfolio is assessed. However, in carrying out this investigation, we repeatedly felt that our work had broader implications for the understanding of the processes of assessment. We thus offer in this paper a case study in the investigation of assessment. We hope that this article will suggest how the issues raised may—suitably interpreted in the light of discipline, level and course design—help to inform and improve assessment practices much more widely.

The use of portfolios

Portfolios have two valuable components for the assessment of professional abilities. First, they contain naturally occurring, authentic evidence of the work of a professional. They thus have at least the potential to be highly valid, offering primary evidence of outcomes achieved rather than secondary evidence such as the ability to talk about how outcomes could be achieved. Second, they involve critical commentary, in which the candidate reflects on the evidence he or she has presented and makes a claim that this evidence shows how they have attained the intended outcomes of the course, and—further—have done so in a way that is demonstrably underpinned by required professional values. However, the assessment of portfolios poses particular challenges, including attaining acceptably reliable assessment.

The course and its assessment

The course studied here, 'Teaching in Higher Education', an accredited Open University 30-point Master's-level course taught by distance learning, has seven learning outcomes. Assessment on the course covers a total of 75 items. There are three broad categories of item:

1. Individual elements, particular outcomes and particular underpinning values, on which assessors form a judgement.
2. Technical requirements, i.e. items whose presence or absence must be recorded, such as a curriculum vitae and a confirming statement from the head of the participant's department.
3. Overall assessments of individual outcomes and then the course as a whole, assessments which are intended to be arrived at by the application of rules to the assessment judgements on individual elements.

These items are listed on the assessment grid in Appendix A. The rules by which assessment judgements on these elements are combined into judgements, first on each outcome and then overall, are shown in Table 1.

The course team has made considerable effort to ensure that the course team, the tutors, the participants, and the assessors share an understanding of what the course is seeking to achieve and how it will be assessed. These efforts include the publication to course participants, tutors and assessors of detailed accounts of learning outcomes, underpinning values and assessment requirements, and showing to course participants illustrative samples from portfolios. Further, in keeping with normal Open University practice, in the run-up to each round of assessments, each
of the assessors independently marks the same one or two portfolios. The assessors then gather together to identify reasons for any differences in the mark they awarded. Through this process, referred to as coordination, they go on to clarify and refine the criteria they will use for that round of assessments.

A third marker, usually a member of the course team, is used to resolve any significant disagreements between the first and second markers. The third marker's assessment stands as the final assessment decision.

It will be seen from Table 1 that there is very little scope for compensation within each outcome and none among outcomes. The course team took the view that, in a course leading to a professional qualification, achieving each outcome was a necessary part of showing professional competence. The impact this has on assessment reliability is considered below.

Reliability in the assessment of portfolios

The face validity of authentic assessments is seldom at issue, but reliability problems are predictably high when validity is high and the achievements in question are complex. Baume and Yorke (2002) investigated inter-rater reliability in the assessment of the portfolios that were used for summative purposes. There were four grading categories on elements of assessment where judgement was required: well achieved; just achieved; not quite achieved; and not achieved. Whilst the assessors tended strongly to agree regarding the fulfilment of technical requirements (as one would expect), the level of agreement was weaker where the assessor's judgement was called into play. The hard criterion of agreement is an exact match between assessors' judgements: here agreement was found on roughly 60% of occasions where elements and values were concerned, and at 39% on course outcomes (where there is a 'snowballing' of the effects of the judgements relating to requirements, elements and values).

Relaxing the agreement criterion to one grade leeway produced agreement percentages in the upper 80s. These figures are broadly consistent with a varied set of studies of inter-rater reliability in portfolio assessment (Herman et al., 1993; Koretz et al., 1993; Nystrand et al., 1993; LeMahieu et al., 1995; Wolfe, 1996; Supovitz et al., 1997; Heller et al., 1998; Pitts et al., 1999). Some technical considerations need to be borne in mind when dealing with the issue of inter-rater reliability. Correlation coefficients, such as the Pearson r, widely used in the literature, are insensitive to level, so it is possible for a high coefficient to arise from two similarly-shaped distributions of grades which are nevertheless differentiated by grade level. Percentage agreement, as has been frequently used in studies, is on its own a poor measure since no allowance is made for chance agreement. The kappa statistic (Davies & Fleiss, 1982), which was used by Pitts et al. (1999) and by Baume and Yorke (2002), is probably less susceptible to distortion than the other two statistics mentioned.

An analysis of the scores from raters of the 53 portfolios used in this study, shown in Table 2, showed clearly that inter-rater variability tended to be greatest where
Table 1. The items assessed on the course, and the scales and combination rules used. A key to the items can be found in Appendix A

<table>
<thead>
<tr>
<th>Items</th>
<th>No.</th>
<th>Contents</th>
<th>How judged?</th>
<th>Grading scale</th>
<th>Rules</th>
</tr>
</thead>
<tbody>
<tr>
<td>0a, 0b and 1a,b to 7a,b,</td>
<td>16</td>
<td>Technicalities; word and page count, the presence of curriculum vitae and of account of teaching</td>
<td>By count or check. Each judgement is independent</td>
<td>Variously: present or not; confirmed or not confirmed; over-length or not.</td>
<td>'Not present' or 'not confirmed' leads to referral or other action as appropriate. Exceeding the word or page limit does not cause a fail. However, assessors are asked to ignore material beyond the limit.</td>
</tr>
<tr>
<td>8a-8e</td>
<td>5</td>
<td>Overall checks including evidence authenticity and age, confirmation of authorship</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1c-g, 2c-f, 3c-d, 4c-e, 5c, 6c-d, 7c-e</td>
<td>20</td>
<td>Particular activities undertaken and requirements met for each outcome</td>
<td>By academic judgement. Each judgement is independent.</td>
<td>Four points; Well Achieved [WA], Just Achieved [JA], Not Quite Achieved [NQ], Not Achieved [NA]</td>
<td></td>
</tr>
<tr>
<td>1c-k to 7c-k</td>
<td>26</td>
<td>Assessments of the six underpinning values, from one to six specified for each outcome</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------------</td>
<td>----</td>
<td>---------------------------------------------------------------------------------------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1-7 Overall</td>
<td>7</td>
<td>Course outcomes</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>By combining judgements on elements (as at right)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Five points; Outstanding Pass [OP], Clear Pass [CP], Bare Pass [BP], Bare Fail [BF], Clear Fail [CF].</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>If any item in an outcome is not achieved [NA], the outcome is failed.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>If just one item in an outcome is not quite achieved [NQ], the outcome should be</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>judged to have been achieved.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>If any outcome is judged a bare or clear fail, the course is failed.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>If 1 to 3 outcomes are failed, resubmission is possible. If more than 3, the course</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>must be retaken.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overall</td>
<td>1</td>
<td>Overall result (pass/fail). By combining judgements on outcomes (as at right)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>By combining judgements on outcomes (as at right)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Two points; Pass, Fail</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>If any outcome is judged a bare or clear fail, the course is failed.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>If 1 to 3 outcomes are failed, resubmission is possible. If more than 3, the course</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>must be retaken.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>75</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 2. Discrepancy rates for underpinning values from assessments of 53 portfolios, set against course outcomes (data from Baume & Yorke, 2002, p. 18)

<table>
<thead>
<tr>
<th>Underpinning value</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 How students learn</td>
<td>5</td>
<td>4</td>
<td>5</td>
<td>5</td>
<td>12</td>
<td>6</td>
<td>6.2</td>
<td></td>
</tr>
<tr>
<td>2 Concern for student development</td>
<td>7</td>
<td>7</td>
<td>5</td>
<td>5</td>
<td>14</td>
<td>7</td>
<td>7.6</td>
<td></td>
</tr>
<tr>
<td>3 Scholarship</td>
<td>5</td>
<td>8</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6.5</td>
<td></td>
</tr>
<tr>
<td>4 Equal opportunities</td>
<td>9</td>
<td>9</td>
<td>11</td>
<td>15</td>
<td>12</td>
<td>11</td>
<td>11.2</td>
<td></td>
</tr>
<tr>
<td>5 Colleagueship</td>
<td>5</td>
<td>6</td>
<td>10</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>7.0</td>
<td></td>
</tr>
<tr>
<td>6 Reflection</td>
<td>12</td>
<td>9</td>
<td>6</td>
<td>10</td>
<td>8</td>
<td>8</td>
<td>8.8</td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>7</td>
<td>8</td>
<td>7.5</td>
<td>6.8</td>
<td>10</td>
<td>7.0</td>
<td>10.7</td>
<td>8.2</td>
</tr>
</tbody>
</table>

Notes: The course outcomes are, in summary: (1) Plan teaching sessions; (2) Teach; (3) Assess student work; (4) Monitor and evaluate teaching; (5) Keep records; (6) Cope; (7) Continue your professional development.

Differences counted in this table are those exceeding one grade.

There are blank cells in the table because not all course outcomes are required to be underpinned by all values.

...
of preparation for the role of assessor of portfolios in which reflective commentary is expected.

The existence of such divergences of view between the assessors, even when considerable guidance, training and coordination had been provided, led us to wonder how assessors were going about making their judgements.

**The research question**

The question that prompted the current study was 'Why and how do assessors make the judgements they make?' The literature on assessment is not very forthcoming about this. It suggests that reliability has been a problem since Edgeworth (1890a, b, quoted in Rowntree, 1996), with Newstead (1996) providing a relatively recent example. The overall message is that the reliability of assessments in higher education is generally not high. Newstead bluntly concludes that assessment in psychology is inadequate, being unreliable, inconsistent, biased, and open to frequent abuse. Psychology is probably no worse than other subjects in this respect, but is at least facing and analysing its difficulties, which is an essential early step towards improvement.

It seemed to us that an understanding of how assessment judgements are made might illuminate other issues in assessment, including reliability and the determination and use of criteria.

**Method**

How might one find out how and why assessors make the judgements they make? Possibilities include the following:

- Sit alongside an assessor as they assess, and ask them to think aloud as they make their judgements. Take notes or audio-record what they say. Perhaps prompt with questions.
- Ask assessors to audio-record their thoughts and reasons as they assess, without another being present.
- Ask assessors to write down their reasons for judgement as they assess.
- Interview or survey them after they have assessed and ask them for their reasons.

Our eventual decision was to ask the assessors to make written comments through on-screen completion of a pro forma, item by item, on which they also recorded their assessment judgements. We did this to maximize the temporal proximity between making the judgements and giving the reasons for judgements, and thus to reduce (though not eliminate) the scope for *post hoc* rationalization.

We decided to undertake an experiment with previously marked portfolios, rather than by studying real-time assessments, on ethical grounds: we wanted to avoid the risk of interfering with the live assessment process.
Table 3. Comparisons between pass/fail judgements of original and experimental assessors, on individual outcomes and overall

<table>
<thead>
<tr>
<th></th>
<th>Course outcomes passed (N = 140)</th>
<th>Overall course passes (N = 20)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No.</td>
<td>%</td>
</tr>
<tr>
<td>Original assessments</td>
<td>136</td>
<td>97</td>
</tr>
<tr>
<td>Experimental assessments</td>
<td>113</td>
<td>81</td>
</tr>
<tr>
<td>Difference</td>
<td>23</td>
<td>-16</td>
</tr>
</tbody>
</table>

Ten assessors, each of whom had previously been briefed and trained to assess on this course, were each paid a fee to assess two portfolios. The assessors had not previously seen the particular portfolios they were asked to assess. Permission was obtained from candidates for this further use of their portfolios.

As well as grading the portfolios, we asked the ‘experimental assessors’ to answer the same question alongside each of the 75 assessment judgements they recorded on each portfolio: ‘What factor(s) persuaded you to make the judgement you made for this element?‘ They made a total of 1500 assessment judgements, including 20 judgements on whether the portfolio represented a pass or a fail. They also provided us with some 29,000 words of reasons and comments.

The experimental assessments

Comparison of scores between original and experimental assessments

The experimental assessments were compared with the original assessments on file. We noted with some surprise that the experimental assessors tended to mark the portfolios lower than was the case with the original final assessments, as is shown in Table 3.

It should be noted that the requirement to pass each outcome in order to pass the course as a whole magnifies the effects of the failure of an individual outcome. Greater weight should therefore be given to the comparisons between judgements at the outcome level than at the overall pass level.

We also list the number of occasions on which each grade was judged to have been achieved. Table 4 shows in finer detail how the experimental assessors awarded

Table 4. Instances of each grade awarded by original and experimental assessors

<table>
<thead>
<tr>
<th>Grade</th>
<th>Clear Fail</th>
<th>Bare Fail</th>
<th>Bare Pass</th>
<th>Clear Pass</th>
<th>Outstanding Pass</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Original (third marker)</td>
<td>3</td>
<td>3</td>
<td>33</td>
<td>67</td>
<td>34</td>
<td>140</td>
</tr>
<tr>
<td>Experimental marker</td>
<td>12</td>
<td>16</td>
<td>41</td>
<td>61</td>
<td>10</td>
<td>140</td>
</tr>
<tr>
<td>Difference</td>
<td>-9</td>
<td>-13</td>
<td>-8</td>
<td>6</td>
<td>24</td>
<td></td>
</tr>
</tbody>
</table>
more lower grades and fewer higher grades than did the original markers. The difference between the two distributions is significant at $p < 0.01$ (Kolmogorov-Smirnov test, two-tailed (Siegel & Castellan, 1988)). Later in the article we explore possible reasons for these differences.

Uncertainties in judgement

As well as asking the experimental assessors why they gave the marks they gave, we also asked them to indicate if, during the process of making each judgement, they had considered giving any alternative judgements and, if so, what these were. The consideration of an alternative judgement suggests an internal debate regarding the making of a judgement. The assessors’ responses are summarized in Table 5, which shows, unsurprisingly, that technical judgements were much less problematic than judgements on values or other non-technical elements.

The consideration of alternative judgements on outcomes is particularly interesting. Assessors are asked to decide a grade on the attainment of each outcome by combining the judgements on each element of the outcome, as described in Table 1. No more than one element may be failed, and that only barely, for an outcome to be passed. This being so, why were assessors considering alternative decisions at the outcome level? In their accounts of their reasons for judgements (considered later in this paper) we sometimes see assessors iterating between their overall impressionistic or holistic judgements on an outcome and the judgements obtained by strict application of the rules. We see them doing this at the level of an individual outcome, sometimes explicitly modifying judgements on an individual element to obtain a result for an outcome with which they feel comfortable. We see a similar iteration between assessment on individual outcomes and the overall pass/fail result. Unfortunately we did not ask them if they considered the alternative option to their overall judgement, on the (as it turns out erroneous) grounds that the ‘decision algorithm’ allows assessors no flexibility of judgement at this final stage.

These findings suggest that some markers are reconceptualizing the assessment process on this course. They do not see the process as making judgements at the level of elements and then applying combination rules to gain an overall result. Rather, they see the process as ensuring consistency, within the combination rules, between elemental and overall assessment.

Reasons given for judgement

The assessors varied considerably in their comments on why they made the judgements they made, ranging from the laconic to the considerably reflective. In this article, attention has been focused on areas of apparent or potential difficulty in assessment, rather than on the seemingly unproblematic areas. Hence the examples quoted are not representative—rather, they illustrate where problems are manifest. In what follows, we have extracted the major themes and issues that we have identified in the comments of the experimental assessors.

Next to each comment from the assessors below, both the grade awarded and any
<table>
<thead>
<tr>
<th>Technical elements (word counts etc.)</th>
<th>Total number of judgements made</th>
<th>Number of judgements where an alternative was considered</th>
<th>% of judgements where an alternative was considered</th>
</tr>
</thead>
<tbody>
<tr>
<td>Values</td>
<td>500</td>
<td>129</td>
<td>26</td>
</tr>
<tr>
<td>Other elements requiring judgement</td>
<td>400</td>
<td>91</td>
<td>23</td>
</tr>
<tr>
<td>Outcomes</td>
<td>140</td>
<td>20</td>
<td>14</td>
</tr>
<tr>
<td>Pass-Fail</td>
<td>249</td>
<td>20</td>
<td>17</td>
</tr>
<tr>
<td>Total</td>
<td>1500</td>
<td>249</td>
<td>17</td>
</tr>
</tbody>
</table>

Assessors were not asked if they considered an alternative to their overall pass-fail decision.
other grade considered are included. After each comment, we have attempted to identify the criterion or rule which seems to us, from the comment, to be have been operated by the assessor in making this judgement—the 'apparent basis of judgement'. In the interests of clarity, we have made very small presentational amendments to the assessors' original comments (some of which were made in note form). We have, in various places, added emphasis to assessors' comments in order to sharpen the focus on the point we are making.

Assessing administrative and technical issues

Assessors are asked to make some administrative or technical checks. A few of these caused the assessors difficulty or uncertainty, and occasioned comment.

Under element 0b (for a full list of course elements, see Appendix A), Account of teaching work undertaken, one assessor wanted qualitative as well as quantitative information about the candidate's teaching, to make better sense of what follows in the portfolio. Under element 8d, Teaching work authenticated by HoD/Course Leader, one assessor commented on the authentication ‘...unless it's a forgery!', with the clear implication that this was very unlikely. Under element 8b, 20% of evidence older than 3 years, one assessor counted to be sure, whereas another deduced the age of evidence from the candidate's curriculum vitae.

On these administrative checks, assessors report adopting a mixed strategy of mostly taking the candidate's accounts on trust, with occasional checks from the evidence in cases of doubt.

Restrictions are placed on word length for claims and numbers of pages of evidence. These restrictions were introduced to manage the workload on assessors. Assessors are asked not to read beyond the word or page limit. Candidates are asked to note word counts and page counts in the portfolio, and most do.

The experimental assessors adopted a variety of approaches to these numerical checks. Some took the candidate at their word, though clearly sometimes with concern about this—for example 'I assume we are to believe the figures candidates give?' Some counted for themselves, describing their counts as accurate or approximate. Some estimated. Some did not say how they made their judgement. Some tried to explain why students might have provided over-length portfolios—'The students here have either forgotten the rules or do not think that you mean what you say about cut-offs'.

Assessors reported considering an alternative judgement about word length and page count in only 4 out of 280 judgements. These were confined to Outcomes 1 and 3, which suggests that any lingering difficulties about making such judgements were resolved as the assessors worked through the portfolios. Minor exceeding of the word limit was sometimes explicitly condoned.

Occasionally, assessors made judgements about the acceptability of content when it did not appear in the expected format. For example, under element 0a, Curriculum vitae, one assessor was willing to accept descriptions included throughout the portfolio in lieu of a curriculum vitae.
The following extended quotation suggests that there is no such thing as a completely unproblematic assessment judgement:

<table>
<thead>
<tr>
<th>Assessment</th>
<th>Alternative considered</th>
<th>Assessor comment</th>
<th>The apparent basis of judgement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Over-length</td>
<td>This candidate is generally too careful to have exceeded the limit carelessly. I have also had to use my judgement to exclude one or two sides where he has clearly included an original document of roughish notes which have been written on the empty side of a piece of scrap paper. If I take such items at face value then I would be finding him guilty of a few more instances of being over length but they are so clearly things written on the back of something else from his scrap paper box that I have exercised my common sense and judgement and have disregarded the relevant sides when I have counted up his pages of evidence - this is another issue for explicit guidance to candidates, perhaps?!? [Portfolio 10, Outcome 7b]</td>
<td>The assessor should be prepared to use their common sense and judgement</td>
<td></td>
</tr>
</tbody>
</table>
The assessment scheme for the course is summarized in Table 1. The rationale for the somewhat unforgiving regime, in which almost all elements and every outcome have to be passed for the course to be passed, was outlined above.

The experimental assessors were additionally asked to record gradings that they had considered but not, in the end, awarded. Occurrences of this are summarized in Table 5. The considered alternative was normally—but surprisingly not always—adjacent to the judgement made.

**Justifications**

Some assessors, responding directly to the task set, have offered what was in effect a justification for the awarded grade. The justification in many instances implicitly reflects criteria:

<table>
<thead>
<tr>
<th>Assessment</th>
<th>Alternative considered</th>
<th>Assessor comment</th>
<th>The apparent basis of judgement</th>
</tr>
</thead>
<tbody>
<tr>
<td>NA</td>
<td>NQ</td>
<td>It is difficult to find any evidence to suggest that the candidate monitors her own work. There is very little to suggest that she is relaying evidence from monitoring into practice. [Portfolio 7, Outcome 4c]</td>
<td>Candidates must provide evidence that they do what they say they do.</td>
</tr>
<tr>
<td>NQ</td>
<td>[... ] It sort of threads the section in a covert way but I just don’t see it addressed head on by the candidate so NQ at maximum, which I will give as the rest is very engaging and interesting to read. [Portfolio 17, Outcome 6d]</td>
<td>A course requirement must be addressed explicitly to achieve a pass on that requirement.</td>
<td></td>
</tr>
<tr>
<td>JA</td>
<td>WA</td>
<td>Evidence includes reflection on tutorials, producing a revision session in response to student requests. The claim however does not say ‘how’ this supports student learning.</td>
<td>A judgement can be lifted to reflect other good material in the portfolio (but not over pass/fail boundary?)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Each course requirement must be met to achieve a pass on the outcome.</td>
</tr>
</tbody>
</table>
I think this is a good and well-focused plan of development, with sections on development for its own sake balanced with those that have a direct pedagogic function and purpose. It’s paced logically over a period of years, shows progression and embraces both subject based and teaching based plans.

A balanced approach is highly valued (the element being seen as ‘well achieved’) by the assessor, as are other qualities underlined at left.

Sometimes an assessor explicitly gave the participant the benefit of the doubt, explicitly allowing their general perception of the person’s ability to affect the particular judgement.

<table>
<thead>
<tr>
<th>Assessment</th>
<th>Alternative considered</th>
<th>Assessor comment</th>
<th>The apparent basis of judgement</th>
</tr>
</thead>
<tbody>
<tr>
<td>JA</td>
<td>NQ</td>
<td>I found this tricky. There is no plan as such, but ways forward are present in some pieces of evidence. On balance, I felt it unjustified to award this outcome an NQ, because this would generate only a BP for the outcome as a whole, and yet much other work here seemed extremely good. A rare case of generosity on my part. [Portfolio 5, Outcome 7e] (In fact, a BP would have made no difference at all to the overall result.)</td>
<td>It doesn’t matter what it is called as long as it is what is needed.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>A judgement regarding a particular element can be lifted to reflect other good material for the outcome</td>
<td></td>
</tr>
</tbody>
</table>

The need to dig for information

Course participants are not allowed to refer in one claim to evidence in another part of the portfolio. This somewhat harsh and unnatural restriction was introduced for
the same reasons as the restrictions described earlier on word length and number of pages—to limit the workload on the assessors. Despite this attempt by the course team to protect the assessors, assessors nonetheless were willing to dig for misplaced evidence in order that they could decide whether the course element had been satisfactorily achieved.

The absence of labels, headings and signposts was noted by some assessors as a contributing factor to their difficulty. In some instances, assessors dug for evidence without success, and graded accordingly.

<table>
<thead>
<tr>
<th>Assessment</th>
<th>Alternative considered</th>
<th>Assessor comment</th>
<th>The apparent basis of judgement</th>
</tr>
</thead>
<tbody>
<tr>
<td>JA</td>
<td>NQ</td>
<td>A difficult mark. The information is available but needs to be hunted -- links to the reflection. But essential information seems to be there. [Portfolio 8, Outcome 5c]</td>
<td>It doesn't matter too much where it is or how it is labelled as long as it's there.</td>
</tr>
<tr>
<td>JA</td>
<td>NA</td>
<td>Any reflection is 'smeared' across the entire commentary. It is a pity that explicit headings have not been provided by the candidate. [Portfolio 2, Outcome 5d]</td>
<td>As with the previous comment - but clear labelling is preferred. (It is unclear why 'not quite achieved' was not considered as a possible judgement here)</td>
</tr>
</tbody>
</table>

**Tension between claim and evidence**

A variant of the difficulty that assessors found in digging for evidence occurred when there seemed to be a tension between what was (or was not) claimed and the evidence put forward:

<table>
<thead>
<tr>
<th>Assessment</th>
<th>Alternative considered</th>
<th>Assessor comment</th>
<th>The apparent basis of judgement</th>
</tr>
</thead>
<tbody>
<tr>
<td>JA</td>
<td>WA</td>
<td>This is a very critical and reflective piece from one end to the other which would justify WA if the claim for this value was more strongly made. [Portfolio 11, Outcome 4k]</td>
<td>To achieve a high grade, a strong claim must be made in addition to the provision of evidence.</td>
</tr>
</tbody>
</table>

**Tension between assessor's grade and comment**

At times the assessor's grade seemed to be contradicted by the associated comment. There were occasional instances when the grade seemed high or low in relation to the comment that was made. Examples include the following:
<table>
<thead>
<tr>
<th>Assessment</th>
<th>Alternative considered</th>
<th>Assessor comment</th>
<th>The apparent basis of judgement</th>
</tr>
</thead>
<tbody>
<tr>
<td>JA</td>
<td>No evidence of reflection, just assertion. [Portfolio 9, Outcome 5d]</td>
<td>Assertion without evidence is just acceptable</td>
<td></td>
</tr>
</tbody>
</table>

Here, the award of JA seems distinctly generous. By contrast, see the next example, where the same grade seems niggardly.

<table>
<thead>
<tr>
<th>Assessment</th>
<th>Alternative considered</th>
<th>Assessor comment</th>
<th>The apparent basis of judgement</th>
</tr>
</thead>
<tbody>
<tr>
<td>JA</td>
<td>Plan is present, it is clear, achievable and links to the evidence. [Portfolio 11, Outcome 7e]</td>
<td>Clear, achievable, evidence-linked plans are just acceptable</td>
<td></td>
</tr>
</tbody>
</table>

*Assessor uncertainty about the judgement*

Occasionally the assessor showed uncertainty regarding the basis of the judgement that had to be made. In the next example, the assessor seems unsure what trade-offs might be allowed.

<table>
<thead>
<tr>
<th>Assessment</th>
<th>Alternative considered</th>
<th>Assessor comment</th>
<th>The apparent basis of judgement</th>
</tr>
</thead>
<tbody>
<tr>
<td>NQ JA</td>
<td>This is a difficult one to grade. The candidate has demonstrated an awareness of equal opportunities, but has not put this in the context of monitoring and evaluating teaching. [Portfolio 7, Outcome 4j]</td>
<td>Awareness of EO (and, by extension, of other values) isn't enough; it must be applied to practice</td>
<td></td>
</tr>
</tbody>
</table>

The same seems to obtain in the next example, but with more than a hint of desperation:
What is happening when we assess?

<table>
<thead>
<tr>
<th>Assessment</th>
<th>Alternative considered</th>
<th>Assessor comment</th>
<th>The apparent basis of judgement</th>
</tr>
</thead>
<tbody>
<tr>
<td>WA</td>
<td>I wonder how to deal with this one in this instance so WA on a wing and a prayer. [Portfolio 17, Outcome 2g]</td>
<td>Unclear.</td>
<td></td>
</tr>
</tbody>
</table>

Criteria perceived as vague or contradictory

<table>
<thead>
<tr>
<th>Assessment</th>
<th>Alternative considered</th>
<th>Assessor comment</th>
<th>The apparent basis of judgement</th>
</tr>
</thead>
<tbody>
<tr>
<td>CP</td>
<td>This [Outcome] is hard to assess as the Portfolio Guide suggests including several things which don’t appear to [fit] the assessment criteria directly. I have concentrated on the assessment criteria. [Portfolio 4, Outcome 5]</td>
<td>In case of a perceived mismatch between advice in different parts of the course material, go with the assessment criteria or go with those which most benefit the students (unclear which). Criterion 6d is particularly vague and would seem to justify almost anything. [Portfolio 6, overall comments]</td>
<td>None apparent.</td>
</tr>
</tbody>
</table>

Sticking to the rules, albeit uncomfortably

Candidates have to do what is asked of them: other material which may denote good teaching may simply not fit the assessment specification, and has to be discarded. Likewise, the assessor is expected to stick the rules, like it or not.

<table>
<thead>
<tr>
<th>Assessment</th>
<th>Alternative considered</th>
<th>Assessor comment</th>
<th>The apparent basis of judgement</th>
</tr>
</thead>
<tbody>
<tr>
<td>CF</td>
<td>cannot substitute evidence not asked for [...] for that which is [...]. The claim must clearly substantiate itself; it is not for the marker to read additional thought into what is written. [Portfolio 11, Outcome 3]</td>
<td>Rules are rules. I must not read into what is given.</td>
<td></td>
</tr>
</tbody>
</table>
Bending the rules?

There were occasions when the mechanism produced a result that the assessor felt did not do justice to the participant. In the following quotation, the whole and the sum of the parts seem not to amount to the same thing.

<table>
<thead>
<tr>
<th>Assessment</th>
<th>Alternative considered</th>
<th>Assessor comment</th>
<th>The apparent basis of judgement</th>
</tr>
</thead>
<tbody>
<tr>
<td>CP</td>
<td>I am a little ambivalent about this and hover between an OP and a CP without bothering about the 'rules' at first. These [the rules] give BP as the only outcome because of three JAs. However the JAs may not be true judgements so with such a good student I want to award CP, so may have to go back and look at the JAs again and jiggle one or two of them.</td>
<td>The assessment judgements must be consistent according to the rules, but it is OK to let a view of a sound overall assessment judgement inform judgements on elements; that is, it is acceptable to iterate between overall and element judgements.</td>
<td></td>
</tr>
</tbody>
</table>

One assessor was concerned that a single element not quite achieved pulled down the overall outcome assessment to a bare pass when the rest of the work suggested a clear pass. In practice, since the course as a whole is assessed on a pass/fail basis, a distinction as fine as this was of no significance to the overall result. However, it clearly mattered to the assessor. In another similar instance, the assessor’s wrestling with the rules was essentially pointless, since the distinction made no difference to the overall result.

<table>
<thead>
<tr>
<th>Assessment</th>
<th>Alternative considered</th>
<th>Assessor comment</th>
<th>The apparent basis of judgement</th>
</tr>
</thead>
<tbody>
<tr>
<td>CP</td>
<td>Overall judgement [of CP] even though it should be BP according to the 'rules'. With 9 areas to consider, two JAs are acceptable I think. Hair splitting I know, but where is 'justice'?</td>
<td>My overall judgement is more important than a slavish adherence to the rules.</td>
<td></td>
</tr>
</tbody>
</table>
Seeking to understand the results

Why did the experimental assessors award lower marks than the original assessors? Possible explanations include:

(1) The additional time taken for giving the comments lowered the mark. Saunders and Davis (1998) have reported this effect in a study of the assessment of undergraduate dissertations, and have suggested that good assessment practice should include common practice on the amount of time an assessor should spend on each candidate's work (in respect of dissertations, they suggest between one and two hours).

(2) An 'experiment effect'. This was known by the assessors not to be a real assessment, which could have affected the grades awarded: however, it not clear why this should produce a consistent shift in one particular direction, as was found here.

(3) Giving reasons for marks lowered the marks given.

The experiment reported here cannot provide an explanation. Some of the experimental assessors had already enquired about the results of the experiment, and been told about the differences in marks from the original assessments. In an attempt to gain some further understanding, each of the experimental assessors was contacted again. They were asked why, in their opinion as an experimental assessor, the marks awarded in the experimental marking were significantly lower than were given by the original assessors. The three possible reasons above were offered, with a request for additional possible explanations. All ten experimental assessors responded. Each of the three possible reasons received some support and some opposition. There were small preferences, on balance, for the views that lower marks resulted from taking more time in the experimental than in the original assessments (because of the time taken to write down their reasons for each judgement) and from the need to give reasons. The suggested 'experiment effect' received less support.

The experimental assessors also made several comments. One assessor felt that taking longer inevitably led to more flaws being found, 'as with the inspection of any piece of work—brick-laying, decorating'. (There was disagreement about whether noting reasons did in fact take longer than normal assessment, as some assessors reported making comments on portfolios as part of their normal practice of assessment.)

Discussing the 'experiment effect' explanation for the lower grades awarded, one assessor commented that they had followed the experimenters' injunction to make the experimental assessment as real as possible, whilst another wondered if, in a real assessment, assessors 'look even harder for grounds to pass people'.

On the 'giving reasons' explanation, one assessor commented:

When you need to justify your thoughts you tend to err on the side of caution. I remember the process of writing reasons to be quite intimidating—this would allow a direct comparison with other [assessors]—due to this I felt I had to be 100% confident of the score.
Another commented on the possibility of the assessor making compromises when 'the marker thinks of the student as a person and ... makes some compromises as a result'. Another is worth quoting in full:

Because we HAD to justify, in detail, everything we did, all conclusions examined, tossed about mentally, we couldn’t HIDE anywhere. Normally, even though the assessment process is fairly rigorous, one can come to an 'overall judgement' without the agonies of decision being made TOTALLY explicit. The more I was forced, in some instances to agonise, the more I began to feel that a lower grade rather than an upper was appropriate.

Further explanations were offered. Unlike the experimental assessors, most real assessors knew their students and had already given them feedback on their work. This, it was suggested, may lead to a bias in favour of the students, a personal bias or a professional bias on the basis of the student’s teaching. It was also suggested that prior knowledge of the student may also lead to a final assessment that measures progress as well as absolute attainment, and the assessor may mentally add such supportive information to that manifested in the portfolio.

One experimental assessor was conscious that their assessment was under very close scrutiny in the experiment, and that absolute adherence to assessment rules was appropriate. One assessor thought that all three explanations applied; another, none of them.

**What criteria emerge from the assessors’ comments?**

Assessors’ comments suggest that they are responding to the assessment rules on a continuum from ‘rules are rules’ through ‘whether I like them or not’ to ‘my judgements are sounder than slavish application of the rules’.

Another continuum stretches from ‘evidence must be clearly and accurately labelled and in the proper place’ to ‘I am prepared to dig for and identify evidence that makes the case’.

There is straight disagreement over whether a claim without supporting evidence should pass or not. There is agreement that, on the assessment of values underpinning practice, both evidence and claim must be present for a clear pass, and also that awareness alone (in this case of EO principles) is not enough for a pass—the principles must be seen informing practice.

Some of these criteria, such as that principles should inform practice, are specified in the course materials and have been reinforced during the moderation process for assessment on the course. Others—perhaps most important and most difficult, the proper relation between judgement as embodied in rules and judgement as an individual and local act—continue to be negotiated, implicitly and explicitly, through the assessment processes of the course. The use of such emergent criteria to improve assessment is considered below.
What do we now know about what is happening when a portfolio is being assessed?

- Even apparently trivial assessment judgements—word counts, the presence or absence of some required element—are not wholly unproblematic.
- Assessors often consider alternatives in making a judgement, even when a coarse (4- or 5-point) scale is used.
- Assessors are generally able and willing to articulate their reasons for the judgements they have made, and do so in a great variety of ways.
- Implicit assessment criteria can be found in these articulated reasons for assessment judgements made.
- Complex and seemingly precise assessment protocols can still leave major residual areas of uncertainty for and disagreement among those using them.
- In an element-based assessment framework, some assessors want to iterate between elements and overall judgements rather than simply to follow rules of combination to achieve an overall decision.
- Asking assessors to give reasons for the marks and grades they award may reduce the grades.
- Two broad causes of difficulty in assessing may be identified. Assessors may or may not properly understand the published outcomes and the criteria. And, largely independent of this, they may or may not agree with the published outcomes and the criteria.

Possible implications for assessment

We note that, by asking assessors to give their reasons for making judgements (and hence by changing the assessment process), we may also change the judgements made. What follows from this?

We might recommend seeking to reduce as far as possible variation in the conduct of assessment. This sits uncomfortably with current assessment practice, in which assessors generally take their own individual approach; as to location, time of day or night for assessing, time taken per script, number of scripts marked in a session, marking papers question by question or script by script, writing comments or not writing on the script being marked, discussing with colleagues particular questions or judgements (or not), making more or less explicit use of any marking schemes or assessment criteria provided, and so on.

We might seek to identify the key variables and their effects. However, as the list above suggests, there are too many variables for an experimental study, even if ethical considerations would permit such an experiment on live assessments (which we believe should not take place without considerable safeguards).

To repeat: we have established so far that differences in the practice of assessment may affect the results awarded. We list above aspects of assessment practice over most of which only a brave course leader would try to exert control, or even influence. Following the work described here, on what potentially significant aspects
of assessment practice might some moves towards an agreed commonality of approach be practicable? Perhaps:

- the explicit use (or not) of marking schemes or assessment criteria
- the recording and justification of results against these
- the provision of feedback to students against the outcomes and assessment criteria.

In considering the implications of these experimental results for the conduct of summative assessment, we cannot avoid asking which set of results we should prefer; those accompanied, or those unaccompanied, by reasons for the judgements made? To the extent that one major function of assessment is to return judgements on the performances of individuals—judgements which will be used to help others form views on the proper future place of the assessed individuals in academe and the wider society—we could argue that we should be wary of any change to assessment practice which may perturb the current distribution of results. It seems that asking assessors to articulate the reasons for their judgements may have such an effect, though we have not proved this conclusively with the work reported here.

However, there are at least two counters to this view. One counter is that a further major function of assessment is to provide a goal for students’ learning, and that clarifying the assessment requirements, albeit in a way that still allows and indeed encourages the individual and serendipitous learning that characterizes education rather than training, will help assessment better to fulfil its role as guide to learning. A second counter is the evidence cited earlier in the paper which suggests that attention really should be given to improving the reliability of assessment. The work reported here points towards a way in which this might be achieved. How?

**Improving assessment**

This work has shown that assessors can in some measure articulate the reasons why they make the assessment judgements they make. There is every reason to expect that this ability could be developed to a more sophisticated level, given time, support and practice. From assessors’ stated reasons for judgement it is possible to infer, with some confidence, as we did above, criteria which may underpin the judgements made. In this extraction of criteria from reasons given we are making explicit what assessors appear to be tacitly valuing in student work. We are thus showing—as the discussion above indicated—areas of agreement and of disagreement among assessors with the published outcomes, criteria and rules.

Going beyond the work reported here, it is possible to share these newly-surfaced criteria and to obtain the reactions of staff, students, employers and other stakeholders. This could in turn lead to the publication and use of refined criteria in teaching, learning and assessment. (It would be important to check that this process did not lead to standards being inadvertently raised, as may have started to happen in this experiment.) A process such as this aligns with Nonaka’s (1994) conception of the production of new knowledge by moving knowledge between the tacit and the explicit domains.
This refining of assessment criteria could contribute to an annual cycle of quality enhancement. More broadly the same general approach could be applied to the way in which a course seeks to construct a learning experience for its students which—as many have commented—is so often led by assessment.

As we noted at the start, this study investigated a course with a very particular, and a particularly complex, assessment scheme. However, we feel that this complexity only magnified issues which may be present, albeit to a lesser and less visible extent, in many assessment processes. In making more visible some of the issues involved in obtaining reliable assessment, the work also suggests a process for improving assessment which has application well beyond the course described.

Acknowledgements

We are grateful to the Open University in several respects. Student Services, via its Director, Professor David Sewart, funded Mantz Yorke’s visiting Professorship at the University for the period during which this work was undertaken, and also the experimental marking of portfolios. The Examinations and Assessment Office and its Director Ben Palmer supplied us with the assessment data in machine-readable form. The Course Team for H851, chaired by Kathy Nicoll, and the Programme Director Dr Barbara Hodgson, both of the OU’s Institute of Educational Technology (IET), supported and encouraged the work. Within the Centre for Higher Education Practice, Eileen Goodyer and Sew Lan Hamilton gave valued secretarial support. Godfrey Pell gave statistical advice, and Carole Baume (the first presentation course team chair for H851) and Professor Graham Gibbs (Director of Research in the Centre) gave academic and moral support. Our discussion of the results has been informed by conversations with Jo Tait of IET and Peter Knight of the Centre for Outcomes-Based Education, again both at the Open University.

The ten experienced assessors who gave us their reasons for the judgements they made—Kemal Ahmet, Rebecca Bartlett, Tony Claydon, Rachel Cragg, Lois Kennedy, Barbara Lee, Jennifer Nisbet, Bob Russell, John Whomsley and Archie Whyte—receive our particular thanks for thus putting their judgement processes under close scrutiny.

Notes on contributors

David Baume, MA FSEDA, is a higher education consultant and an elected member of ILTIE Council and Accreditation Committee. Previously he was a Director of the Centre for Higher Education Practice at The Open University. David was founding chair of the Staff and Educational Development Association and founding co-editor of the International Journal for Academic Development.

Mantz Yorke spent nine years in schools and four in teacher education before moving into staff development and educational research. He then spent six years as a senior manager at Liverpool Polytechnic, after which he spent two
years on secondment as Director of Quality Enhancement at the Higher Education Quality Council. On return to his institution he has researched and reflected on various aspects of the student experience.

Martin Coffey, BSc MSc Chartered Occupational Psychologist, is currently pursuing a dual role; as an independent business psychologist as well as holding a 0.5 lecturer post at the University of Leicester, where he teaches on distance learning MSc programmes in Occupational Psychology. Previously he was a Research Fellow in the Centre for Higher Education Practice at The Open University. His current research interests are in the area of teamwork in academic settings.

Note
1. We use the terms ‘experiment’ and ‘experimental’ in a broad sense in this article.

References
Appendix A. The elements of assessment

0. Introduction
0a Curriculum vitae
0b Account of teaching work undertaken

1. Design of teaching sessions
1a Up to 750 word claim
1b Not more than 30pp of evidence
   Teaching methods appropriate to:
1c Learning outcomes
1d Learners
1e Mode of study
1f Subject
1g Resources
   Values:
1h How students learn
1i Concern for student development
1j Equal opportunities
Objective 1 overall

2. Teaching and learning methods
2a Up to 750 word claim
2b Not more than 30pp of evidence
   1st teaching and learning method
2c Observed and reported
2d Use of appropriate learning technology.
   2nd teaching and learning method
2e Observed and reported
2f Use of appropriate learning technology.
   Values:
2g How students learn
2h Concerns for student development
2i Equal opportunities
2j Reflection
Objective 2 overall

3. Marking and feedback
3a Up to 750 word claim
3b Not more than 30pp of evidence
   Feedback and marking
3c Feedback to students
3d Grading and marking
3e How students learn
3f Concern for student development
476  D. Baume et al.

3g Equal opportunities
3h Reflection
Objective 3 overall

4. Monitoring and evaluating teaching
4a Up to 750 word claim
4b Not more than 30pp of evidence
Evaluation
4c Self
4d Peer
4e Student
Values:
4f How students learn
4g Concern for student development
4h Scholarship
4i Colleagueship
4j Equal opportunities
4k Reflection
Objective 4 overall

5. Support and administration
5a Up to 500 word claim
5b Not more than 20pp of evidence
Records:
5c Appropriate records kept
Values:
5d Reflection
Objective 5 overall

6. Coping strategies
6a Up to 500 word claim
6b Not more than 20pp of evidence
Resource management
6c Time management
6d Operating within resources
6e Colleagueship
6f Reflection
Objective 6 overall

7. Reflection and planning CPD
7a Up to 500 word claim
7b Not more than 20pp of evidence
7c Reflection
7d Needs analysis
7e Continuing Professional Development Plan
Values:
7f How students learn
7g Concern for student development
7h Scholarship
7i Colleagueship
7j Equal opportunities
7k Reflection
Objective 7 overall
8. Overall checks
8a No one outcome wholly by simulation?
8b < 20% of evidence older than 3 years?
8c 'Confirmation of authorship' signed?
8d Teaching work authenticated by Head of Department/Course Leader?
8e Have you tutored this student?

9. Overall Assessment
Pass or fail; resubmission is possible with up to three outcomes failed, a re-retake is required for more than three outcomes failed
Appendix 1 – Citations, Reviews and Refereeing

Citations of Baume 2001:2
"Baume (2001) also argues that where a course has multiple learning outcomes, reliability almost inevitably fails. Therefore to increase the reliability of the overall assessment, the number of learning outcomes needs to be reduced."

"Some of the models, most notably the toast rack, include no overall reflection on or critique of the learning that has occurred – which some educators regard as the process through which real learning takes place [ref to Baume 2001:2 at this point]"

Citations of Baume and Yorke (2002)
"The level of chance agreement will depend on the distribution of scores given by each assessor, which in turn depends on how well the assessors performed. We believe the method some authors use to calculate the 'theoretical distribution' of agreement is inappropriate (Baume and Yorke, 2001)."

Eight references to Baume and Yorke (2002), including:
"Different studies find widely different rates of inter-rater agreement (see Baume and Yorke 2002)."
"(4) Areas where it is hard to get agreement
"Particular areas in assessment provoke more disagreement than others (Centra 1994: Baume and Yorke 2002), suggesting differences in perspective about what is to be valued in a portfolio. These areas tend to be, hardly surprisingly, less factual areas and areas where inadequate contextual information is available. In general, clear definitions of portfolio assignments, greater discussion among assessors, and ensuring that all
assessors have equal access to relevant information among assessors are advocated as means of increasing agreement on particular items. Baume and Yorke report that the Open University course team 'now emphasises previously problematic elements in briefings for assessors (i.e. improved training) and course tutors (i.e. clearer assessment tasks)'. If this emphasis includes discussions about the views of the assessors, the inclusion of their perspectives in actions taken and the reaching of a consensus about what is reasonable and to be valued, this would be close to an interpretivist view...

"(5) Holistic versus individual element scoring

"Positivist researchers discuss whether it is preferable to rate the portfolios holistically or by task across all portfolios (Nystrand et al., 1993, Pitts et al 1999; Baume and Yorke 2002)..."

"Although some of the research is carefully done and well worked out, producing useful insights (e.g. Supovitz et al. 1997, Baume and Yorke 2002), many researchers in this tradition seem unaware of the theoretical and philosophical framework in which their work is rooted...."

"In principle, researchers in the different assessment traditions have fundamentally different ways of conceptualizing and researching assessment. In practice, divisions are less stark. For example, many who operate in the conventional, positivist mode of assessment are aware of the dangers of achieving 'reliability' through measuring only trivial aspects of work. For example, Baume and Yorke (2002) describe the care taken in an Open University course for accrediting university lecturers to assess a course 'validly' by constructing clear guidelines for assessors which reflect course aims and practice..."


"The potential for staff/student coherence in understanding of the assessment task is further increased when the students are provided with examples of the criteria in use, such as when previously assessed material is made available. Baume and Yorke (2002) give, in passing, an example of this practice."

_Citation of Baume and Baume (1996:1)_


"A third factor associated with the motivating potential of a learning setting is the type and timing of the feedback provided. Students report higher
motivation when feedback is relevant and timely (Baume and Baume, 1996)."

Citation of Baume and Baume (1996:4)

"...Such portfolios are used widely in programmes for the education and accredit of teachers in higher education. The UK Staff and Educational Development Association (SEDA) has since 1992 recognised some 60 such programmes in HEIs, in UK, Australia, New Zealand, Hong Kong, Singapore and Sri Lanka, and has accredited some 1700 teachers who have successfully completed recognised programmes (Baume and Baume, 1996)..."

Citation of Baume (1996:2)

"...However, despite these advances, Baume correctly observes that ‘the very concept of a scholarship of pedagogy is still very unfamiliar to university teachers.’" (Baume 1996 p 4)

Review of Baume (1996)

"...David Baume sets out to marry the notion of autonomous learning with students in a higher education setting – a difficult proposition given the two concepts of autonomy and formal education. The formal education setting has a culture of being prescriptive in what is taught and how it is taught. David Baume challenges this historically taken-for-granted system, one which excludes the student experience in curriculum design and delivery to a greater or lesser extent. He outlines what could be achieved through the development of Learner Autonomy, what needs to be done, the advantages as well as disadvantages and, more importantly, why change is needed. As an adult educator I empathize with those needs, particularly in higher education..."

Review of Kahn and Baume (2003)
“This is a great book. I started flicking through this the way you do when you look at an edited collection and I found myself drawn in time and time again to the text. Of course, I recognise a lot of the names of the chapter authors and I felt I understood many of the key issues, but I actually found it difficult to put down because the authors reframe a number of key concepts with which I am familiar and throw new light upon established questions and issues.

“Just about every chapter has something to offer, even to an old hand like me, so it is invidious to pick out individual chapters for comment...

“...This book will be really valuable both to people just starting out in this domain and also to those who have been immersed in it for decades. “

Referees’ comments on Baume and Yorke (2004)

Referee 1

“An interesting and important contribution; methods are appropriate to questions asked. Findings are important, for example that making comments / using criteria might reduce the mark awarded. There are a few minor typos…”

Referee 2

“Opens up thinking about the cognition processes of development in a way that had not been done before. Very important. A day or so before I received the paper I had read Baume and Yorke (2002). It seems to me that this paper is probably by the same authors (and about the same group of assessors). If so then the links need to be made in this paper. If this is a second round - say so. This is the first paper I’ve read that I have had no qualms about.”

Citation of The International Journal for Academic Development


“...I would like to end by paying tribute to my founding co-editors, Patricia Weeks and David Baume, both of whom have now retired from the journal. David Baume in particular was the true inspiration behind IJAD, the person who saw a need for such a journal and convinced our publisher that such an idea could be viable and even modestly profitable. David’s foresight and indefatigable energy as editor helped to provide a forum for educational developers world-wide, and also contributed in no small part to the building of ICED, for which IJAD represents its best known and important activity. Educational developers across the world owe David a sincere debt of gratitude. ...”

David Baume
Appendix 2 - Co-worker statements

Carole Baume, Publications 1 and 7
The Open University

Professor Alan Jenkins, Publication 2
Oxford Brookes University

Professor Mantz Yorke, Publications 8, 9 and 10
Liverpool John Moores University
PhD by Publication – David Baume


The paper describes the development of the initial SEDA Teacher Accreditation Scheme, its take-up, its quality assurance framework and further developments for specialist groups of learning supporters.

Like most successful development work, the writing of the paper and the development of the SEDA Teacher Accreditation Scheme that the paper describes were very much collaborative efforts. I shall detail here, as far as I can, the particular contributions of David to the paper and the scheme.

I was a founding member of a SEDA working group which developed the original ideas for the Teacher Accreditation Scheme. To write the paper, David and I started from a collection of committee papers, development notes, scheme leaflets and notes of research interviews with scheme participants, both programme leaders and new lecturers taking the programmes. We then spent time identifying for the paper the most important themes and issues in the development and implementation of the scheme. This involved us in standing back from the detail to see the overall picture. An early draft of the paper provided the basis for a workshop that we ran at the first conference of the International Conference for Educational Development held in Vasa, Finland in 1996. We gained many ideas and helpful feedback from participants in this workshop. We researched and added material about current and future developments of the scheme, and completed the final paper for IJAD 1:2. We each drafted sections of the paper. David brought these together into the final manuscript. We revised this for publication together. David played the major role in writing the conclusion and discussion.

I saw David take two complementary roles in the development of the Teacher Accreditation Scheme. First, he worked as a technical advisor to the SEDA working group. The group’s initial idea had been to devise a London-wide programme to train new higher education teachers. We soon saw that such a programme would not be practically possible, given travel issues and also organizational practicalities. David encouraged us, instead of giving up, to ask what would be the learning outcomes of such a programme had it been possible to run it. David and I then built on our experience of together developing an accreditation framework for playworkers, helping the SEDA working group to make the outcomes clear and simple. The working group enthusiastically adopted the idea that David had developed for the playworkers scheme that specified values should underpin teaching, as any professional work.

Secondly, David was Chair of SEDA at the time. In this role he was a very active sponsor for the development of the Scheme. He led and supported SEDA to work out
what to do with the statement of outcomes and underpinning values that the working group had produced. He had a major input into the form of the Teachers Accreditation Scheme, which ran with only minor changes for ten years. There were two main elements to the Scheme. A statement of course outcomes and values; and SEDA recognition – through what participants told us was a formal, rigorous and supportive process – of programmes that could show that, first, they developed staff towards these outcomes and values and, second, that programme participants had to demonstrate attainment of these outcomes and values to pass the programme and thus gain SEDA accreditation.

David provided similar input to the development of the Fellowships Scheme, with me chairing the Fellowships Committee, and also to the development of an early form of what is now the Professional Development Framework. David also took a major role in the EduLib project to develop a version of the Teacher Accreditation Scheme for Library staff who teach.

Carole Baume BSc MA FSEDA
Regional Director, The Open University in the North West
23 September 2004
PhD by Publication – David Baume


The chapter is based on research into current evaluation practice by developers, which we jointly designed and carried out, supported by a small grant from the Staff and Educational Development Association.

David drafted the adaptation of Nevo and the worked example, and we worked on subsequent drafts of the chapter together.

Carole Baume BSc MA FSEDA
Regional Director, The Open University in the North West
23 September 2004

Co-worker: Professor Alan Jenkins: then a member of the Oxford Centre for Staff and Learning Development.

This evaluation is a critical appraisal of an educational innovation primarily / immediately to guide developments within the University –and then to inform developments outside Oxford Brookes.

To explain the context: I developed the idea of a whole range of educational development events over one term in 1996 at Brookes with respect to the educational use of information technology. The intent of the term was both to raise awareness but also to ensure critical appraisal to shape practice and policy – to repeat mainly at Brookes but also elsewhere. For the immediate context was an innovative educational institution that had made limited investment in information technology but now needed to decide whether and how to proceed. During IT Term there was a profusion and variety of events we in the central educational development unit organised, but so did all 12 academic Schools, the library, the student union and, yes, the university computing centre.

David was brought in just before the term began and an outline evaluation framework that he proposed was agreed.

So we wanted the impact rigorously but quickly evaluated – for we wanted to end the term with a clear set of ideas and policies to take the institution forward – somewhere! We also wanted an evaluation that would inform very different ‘stakeholders’ from central management to staff in particular disciplines. The evaluation was by definition seeking to capture and analyse something very complex, fast moving and very multifaceted. Those of us centrally involved in running the events had limited time to guide David – and anyway we knew that for the evaluation to have any credibility we needed to keep some critical distance away.
The evaluation was quite brilliantly designed; capturing through document study, observation, interviewing ... the complexity as seen from the angles of the various stakeholders. The final report was delivered quickly and so timely in that we could then use it to shape policy – a central requirement. It was directed to those shaping overall institutional policy, those in the Schools, the educational development unit ... Each of us could see what we had done clearly portrayed and then clear evidence based suggestions to take us forward.

The great danger of such an event is that the events are all consuming – but evanescent. The evaluation that David designed analysed them carefully and with insight; rigorously but respecting the ‘messyness’ of the events – and then painted pictures and insights to take us forward – and these insights came from the data but also set within a rigorous educational framework. We then mined the publication to also inform others outside Brookes of the term and what we had learned. It was and is an exemplary evaluation and publication.

Professor Alan Jenkins
Westminster Institute
Oxford Brookes University
September 27, 2004
Co-worker statement in respect of PhD submission by David Baume

During the academic year 1999-2000 I was Visiting Professor at the Centre for Higher Education Practice [CeHEP] at the Open University. My role included working with colleagues in CeHEP, including David Baume, to develop their research and publication profile. It quickly became apparent that the courses run by CeHEP at that time for teachers in higher education had generated a lot of data which had the potential to lead to useful research findings. The three co-authored items discussed below derive from studies based on the CeHEP provision.


In recent years there has been a growth in interest in so-called 'authentic assessment', in which the use of portfolios has become prominent. The 'Teaching in Higher Education' course run by David Baume at CeHEP required students to fulfil a complex set of learning outcomes, with this being captured in a portfolio assessment involving 75 judgements of varying kinds. Each portfolio was marked by two assessors, with a third marker available to resolve any significant boundary disputes.

The literature on portfolio assessment suggested that the reliability of assessor judgements was problematic. The question for the CeHEP course was whether such a tightly specified assessment regime would suffer from problems with reliability as much as looser approaches that had typically been used in US school education. The OU archives were able to provide us with data from 53 portfolios, each double-marked, for analysis. The research into the reliability of the CeHEP assessments was a genuinely joint affair, with extensive discussions between us regarding the best way of handling the dataset, the indexing of the reliability of the assessments, and the way in which the implications for practice might be taken forward. Of course, each of us took a leading role at different times as the study developed.

At first blush, an analysis of 53 portfolios might be seen as a relatively small-scale affair. However, the analysis of 53 portfolios x 2 markers x 75 separate judgements involved sustained attention before the results could be 'boiled down' into a form that would be useful for others. The pinpointing of areas of comparative unreliability indicated where curriculum improvements might most usefully be made, and the research offers a template for others interested in reliability in assessment, whether or not in respect of portfolios.

Whereas Baume and Yorke (2002) drew on archival data, this article required the designing and execution of what can, broadly, be termed an experiment. The study was experimental in that it involved new empirical work intended to probe more deeply into the question of reliability of portfolio assessments. In this sense it is a successor to the previous study.

David was able to retrieve, again from the OU archive, ten portfolios that had already been assessed, and for which we had the original assessments. The experimentation lay in David getting new, but experienced, assessors to re-assess the portfolios and to provide a record of their thinking as they made their judgements. This provided both statistical and narrative detail, the latter being extensive. Martin Coffey was involved in collating the qualitative findings, and in discussion of the results.

The main findings have raised interesting questions about the assessment process. In brief, the 'experimental assessors' marked more toughly than the original assessors: was that perhaps a consequence of their feeling obliged to pay more exacting attention to the expected learning outcomes? At times the narratives suggested that this might be the case, whereas at other times they suggested that assessors were bending the 'rules' in order to come up with a result with which they felt comfortable.

The findings of this study are disturbing for assessment practice, since they imply that the judgements commonly reached regarding students' performance may be less robust than they are typically portrayed. This is a study for which there is scant prior literature, and in my judgement it breaks new ground.


I was invited to write a chapter for this book, but felt that the work in Baume and Yorke (2002) would provide a more useful case study than anything that I had to offer as a solo author. The thrust of the book was aligned more with David's interests than mine, and it seemed sensible to pass the opportunity to him to lead on the writing, with me acting in very much a minor, supporting role. David drafted the chapter, and after discussion we agreed the final version.

This case study took the formal research of the 2002 article and converted it into a format more likely to suit the intended readership of the book - educational developers and academics who are not necessarily educational researchers. It may, as a result, have more influence on practice than the original research paper (which, of course, is available for anyone wishing to follow it up).

**Overall**

These three pieces reflect the collaboration - and at times tough but constructive debate - that took place during the formal research process, and also the fact that we were able to reach a principled agreement at the end of it from positions that were not always aligned.
The research has influenced David's subsequent thinking about frameworks for accreditation. (I doubt that any new framework will demand 75 separate assessment components, following our studies!)

Universities UK, SCOP and HEFCE have recently sponsored work on the measurement and recording of student achievement: our researches should give pause for thought to those who have a commitment to honours degree classifications and grade-point averages, with their implicit confidence in the precision of such 'measures'.

Mantz Yorke
Professor of Higher Education
### Appendix 3 – Summary of work on accreditation

<table>
<thead>
<tr>
<th>Accreditation Scheme / Client</th>
<th>Dates</th>
<th>Activities / role</th>
<th>Extent of contribution</th>
<th>Comments</th>
</tr>
</thead>
</table>
| Playworkers through the National Children’s Play and Recreation Unit | 1989-90 | A Devising playwork accreditation framework and process  
B Planning implementation  
C Writing scheme documentation and manuals  
D Selecting and training assessors | Lead contributor at activities A and C and joint lead at B and D | The scheme went into operation. Clear traces of the original scheme, particularly the use of underpinning values, are still evident in current NVQs in playwork |
| Institution of Environmental Health Officers (IEHO)  
‘Assessment of Professional Competence’ scheme | 1990-91 | A Designing assessment process  
B Writing candidate and assessor manuals  
C Designing and running training assessors of first assessors | Lead contributor at each stage | The current scheme documentation is strongly based on, and sometimes very similar to, that which I prepared. Some 3000 EHOs have been through the process |
| Staff and Educational Development Association (SEDA)  
Teacher Accreditation Scheme | 1990-2003 | A As Chair of SEDA, strongly supporting and encouraging the development at every stage  
B Contributing much conceptual and development work on the accreditation framework | Leading, facilitating, supporting, providing substantial technical input | By 2003 when the scheme was closed to new applicants (deferring to the ILTHE) some 3100 teachers had been accredited and some 65 programmes recognised in 6 countries |
<table>
<thead>
<tr>
<th>SEDA Fellowship (for staff and educational developers)</th>
<th>1994-2004</th>
<th>A. As Chair of SEDA, strongly supporting and encouraging development at every stage B. Collaborating on development of the scheme C. Mentoring and assessing candidates</th>
<th>Supporting, substantial technical input</th>
<th>In October 2003 44 Fellowships and Associate Fellowships had been gained and a further 29 candidates were working towards the qualification</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dearing Committee</td>
<td>1996</td>
<td>Acting as the Coordinator of SEDA’s Accreditation activities</td>
<td>Leading the production of SEDA’s strong recommendation (written and oral) to Dearing that higher education teacher accreditation should be introduced</td>
<td>This contribution, along with a similar view from AUT, led to Dearing’s recommendation and the setting up of ILTHE and later The Higher Education Academy</td>
</tr>
<tr>
<td>The Booth Committee</td>
<td>1997</td>
<td>A. As SEDA Accreditation Coordinator, working with AUT to establish and fund the Booth Committee on the Implementation of Dearing’s recommendation on teacher accreditation B. Member of working group undertaking detailed standards development</td>
<td>A. Leading the SEDA team B. Much drafting work for the Committee on the detailed standards</td>
<td>The standards proposed by Booth had a clear and direct impact on the current ILTHE standards. ILTHE has accredited (Nov 2003) 147 programmes and admitted some 17,000 members</td>
</tr>
<tr>
<td>Biotechnology and Biological Sciences Research Council (BBSRC 2001) – Research Supervisor Accreditation programme</td>
<td>1998-2000</td>
<td>A. Devising and consulting on accreditation framework and process B. Training first round of assessors</td>
<td>Project co-leader with a colleague from BBSRC</td>
<td>The BBSRC scheme is now in operation</td>
</tr>
<tr>
<td>Institute for Learning and Teaching in Higher Education (ILTHE) / SEDA</td>
<td><strong>2000</strong></td>
<td>As leader of SEDA Panel, successfully proposing that ILTHE accredit all SEDA Teacher Accreditation courses <em>en bloc</em> as giving ILTHE membership</td>
<td>Leading the SEDA team, preparing documentation and making the case at the accreditation event</td>
<td>At November 2003, 39 of the 147 ILTHE-accredited programmes had been accredited via their prior SEDA accreditation</td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>ILTHE</td>
<td><strong>2002-04</strong></td>
<td>Elected member of Council</td>
<td>Member</td>
<td>Overall ILTHE policy and monitoring</td>
</tr>
<tr>
<td></td>
<td><strong>2002-04</strong></td>
<td>Member of Accreditation Committee</td>
<td>Member</td>
<td>Helping set and review policy and practice</td>
</tr>
<tr>
<td></td>
<td><strong>2002</strong></td>
<td>Trained and qualified as accreditor, reviewing initial entry applications</td>
<td>Operational role</td>
<td>Updating my expertise in accreditation</td>
</tr>
<tr>
<td></td>
<td><strong>2003-04</strong></td>
<td>Drafting Code of Practice for Lecturers at the request of ILTHE Council</td>
<td>Chairing the working group</td>
<td>Developing a Code of Practice is a further stage in the professionalisation and accreditation of lecturers</td>
</tr>
<tr>
<td>SEDA Professional Development Framework (PDF)</td>
<td><strong>2001-04</strong></td>
<td>A Devising accreditation framework and process B Planning implementation C Writing scheme documentation D Mentoring programme leaders E Devising a specialist qualification within SEDA-PDF</td>
<td>As a SEDA Accreditation Coordinator, jointly leading the development of SEDA-PDF. SEDA-PDF (SEDA 2004) requires explicit attention to the process of development as well as to specific and specialist capabilities and outcomes</td>
<td>SEDA-PDF was planned to meet continuing demand from institutions for a SEDA framework, despite the establishment and wide take-up of ILTHE accreditation. Since its launch in 2002, 20 institutions have registered with the scheme of which 13 have so far been recognised and 25 programmes registered of which 17 so far recognised</td>
</tr>
</tbody>
</table>

David Baume
<table>
<thead>
<tr>
<th>Project/Role</th>
<th>Year(s)</th>
<th>Description</th>
<th>Responsibility</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Associate Lecturer Development and Accreditation Pathway (ALDAP) for Open University tutors</td>
<td>2000-03</td>
<td>A Devising accreditation framework and process B Planning implementation C Writing scheme documentation D Training accreditation coordinators E Planning for external (SEDA) Accreditation</td>
<td>Lead project consultant on all aspects. ALDAP matches the SEDA-PDF framework to the particular roles of an OU tutor. It also provides a smooth transition between initial and continuing professional development</td>
<td>ALDAP is available to 7500 OU tutors. Some 360 had enrolled by mid-January 2004</td>
</tr>
<tr>
<td>National Health Service University (NHSU)</td>
<td>2003-04</td>
<td>Drafting standards for NHSU supporters of learning (tutors)</td>
<td>Consultant</td>
<td>The first phase of this work is complete</td>
</tr>
<tr>
<td>DfES New HE Teaching Standard</td>
<td>2004</td>
<td>Working group membership</td>
<td>Participation in meeting. Providing detailed feedback on draft</td>
<td>Builds on all previous work, with the intention to focus on the disciplines; on values, knowledge and competence; and on smooth transition from initial to continuing professional development</td>
</tr>
<tr>
<td>The Higher Education Academy</td>
<td>2004</td>
<td>Consultancy on professional standards development</td>
<td>Reviewing a wide range of professional standards to determine appropriate and useful features for a new standard. Planning and facilitating consultation events towards the new standards</td>
<td>This review is complete and with The Academy. It and the events will inform the development of the new HE teaching standard</td>
</tr>
</tbody>
</table>