e is for exploration: Assessing hard-to-measure learning outcomes

Kerry Shephard

Kerry Shephard is a professor of Higher Education Development in the Higher Education Development Centre at the University of Otago. Address for correspondence: Professor Kerry Shephard, Higher Education Development Centre, University of Otago, Dunedin, New Zealand. Email: kerry.shephard@otago.ac.nz

Abstract
The focus of this paper is on the use of e-technologies to enable higher education to better assess aspects of learning that have proved difficult to assess using more conventional means. Higher education describes the knowledge and abilities it intends its graduates to have acquired before graduation, and it has a wide range of approaches to assess these. Higher education also seeks affective outcomes in the form of values, attitudes, behaviours and related attributes or dispositions, and these have consistently proved more difficult to assess by examination or assignment. After graduation, however, graduates are often assessed within the professions via portfolios, interview and peer or expert review. Assessment may focus on teamwork and networking skills, productivity, creativity and values fit to the profession. How can e-technologies help with these forms of assessment? This paper reviews some of the e-based approaches and explorations that have supported or could support assessment of affective attributes. At each stage, the paper seeks to establish the common elements of assessment in the different regimes and how e-assessment contributes, or may contribute in the future. The paper concludes that many of the strengths of e-assessment lend themselves to an evaluation paradigm rather than to conventional assessment for intended learning outcomes.

Introduction
The key focus of this paper is not on what e (as in electronic) can do, but on what e may be able to do better, or differently, from traditional approaches. This paper sets out to explore the more challenging aspects of assessment in higher education and, on the way, attempts to identify how e-technologies are being used, or could be used, to meet these challenges, on either side of graduation. Exploration of this order needs to be broadly based. It is necessary, for example, to explore some of the remarkable developments in the use of portfolios in education, alongside any investigation of how e-portfolios may operate. This is attempted in the subsequent paragraphs.

© 2009 The Author. Journal compilation © 2009 Becta, Published by Blackwell Publishing, 9600 Garsington Road, Oxford OX4 2DQ, UK and 350 Main Street, Malden, MA 02148, USA.
Higher education attempts to describe the knowledge and abilities it intends its graduates to have acquired before graduation, primarily via intended learning outcomes, and it has a wide range of traditional and innovative approaches to assess these in a valid, reliable and workable manner. An important element of most forms of assessment in higher education is the use of assessment criteria to select individuals for particular descriptive categories, such as pass, merit, class and honours. Examinations and assignments have provided the traditional mainstay of assessment, but a wide range of additional approaches have contributed, perhaps particularly in recent years. Assessment has become a major focus for educational research and development. Higher education also seeks affective outcomes in the form of values, attitudes, behaviours and related attributes or dispositions, some described as graduate attributes, and these have consistently proved more difficult to assess by examination, assignment or other traditional approaches (described recently by Barrie, 2004; James & Brown, 2005). Often, these ‘other’ outcomes are not openly assessed and they remain indicative or aspirational on behalf of institutions (Carter, 1985; Jickling, 2003; Shephard, 2007, 2008).

Assessment and selection after graduation, however, for appointment or promotion, use processes in ways often markedly different from those that characterise assessment in the academy. Graduates are often assessed within the professions via portfolios, interview and peer or expert review. Assessment in employment often focuses on teamwork and networking skills, productivity, creativity, initiative and values fit to the profession rather than on disciplinary knowledge, and in that way perhaps has more in common with assessment for affective attributes than for cognitive ones. In recent years, higher education has begun to explore how outcomes other than knowledge and cognitive skills can be assessed. Bryan and Clegg (2006) provide a recent overview of the diversity of assessment in higher education.

Alongside these domains of development, all branches of society have been exploring how information and communication technologies (ICT) can support and extend conventional activities, including these assessment and selection processes on either side of graduation. We know that in generic terms, ICT can add flexibility (allowing adaptation and exploration), interactivity (allowing links between different elements of the same work and to different works in multiple media) and connectivity (allowing interactions between multiple players). This paper focuses on aspects of this exploration of ICT that combine enquiry into learning outcomes that traditional approaches find difficult to determine, with varying approaches to assessment and selection undertaken before and after graduation. Throughout, an attempt is made to enable learning in each domain to benefit application in the others. Can, for example, the development of social networking technologies in society help to support assessment processes in higher education? And how can the exploration of e-portfolios in higher education contribute to their wider and appropriate use in the professions? Because the author’s experience has been predominantly within higher education, and because higher education (as an entity particularly involved in assessment) would be expected to display good practice in assessment, this paper focuses on either side of graduation, within this context.
What aspects of learning prove difficult to assess?

It may be that other authors in this series will attempt to define assessment, electronic or otherwise. This paper keeps its options open and uses broad descriptions and, where appropriate, exemplars to convey something about the nature of assessment, its association with selection and its underlying characteristics. This section seeks to describe some of the elements of learning that traditional approaches to assessment have found difficult. James and Brown describe a range of enduring issues surrounding the improvement of learning outcomes. The range includes ‘the difficulty of clearly conceptualizing some aspects of learning that are seen as highly desirable (e.g., attitudes, dispositions, values, identities), but do not have a common interpretation in the way that straightforward practical or cognitive skills do’ (James & Brown, 2005, p. 9).

Such conceptualisation may be difficult, but this does not prevent many institutions from seeking affective outcomes in the values and beliefs that its graduates hold to ensure that they are consistent with a particular view on their future role in society. The University of Sydney provides a particularly straightforward example:

University of Sydney, Graduates of the Faculty of Veterinary Science will hold personal values and beliefs consistent with their role as responsible members of local, national, international and professional communities (e.g., protect the natural environment, maintain biodiversity and conserve endangered species) (http://www.vetsci.usyd.edu.au/future_students/undergraduate/graduate_attributes.shtml [accessed August 26, 2008]).

Many universities identify the attributes that they hope their graduates will aspire to, but (and perhaps particularly where these attributes relate to affective characteristics) these attributes are rarely assessed with the same level of objectivity as those described by intended learning outcomes.

There are exceptions. Shephard (2008) has reviewed affective teaching in a range of subject areas in higher education. Some programmes are clearly designed to change students’ affective attributes to suit the needs of a particular occupation or profession. Shephard describes how medicine, for example, adopts particular approaches to teach and to assess. Health care professionals learn about healing; they learn the skills to heal; and often their education and assessment attempt to ensure that they have caring attitudes towards patients (see, e.g., Howe, 2002). Objective Structured Clinical Examinations are now widely and summatively used in clinical education, and these are designed to specifically assess the attainment of some affective learning outcomes. There are also, however, examples in the literature where the validity of assessments in the affective domain have been called into doubt (see, eg, Gratton, 1996).

Universities are also under pressure from external stakeholders. Much research demonstrates that employers seek attributes in our graduates that do not lend themselves to assessment against intended learning outcomes. One recent survey in Australia reported that ‘Creativity and flair was the [academic] skill which was considered to be most important to employers overall but … performance of graduates is poorer than its
importance would warrant ...’ (DETYA, 2000, p. 36). Along similar lines, but in the UK, the Roberts Report (2002) recommended that research students should have 2 weeks of transferable skills training each year, based in part on feedback from the private sector that postgraduates were often highly competent in technical expertise but lacked broader skills such as teamwork. The UK Research Councils produced a joint statement on skills, listing 36 critical skills (RCUK, 2001). Many of these are affective attributes, including a willingness to learn, open-mindedness, self-assessment, self-reliance and self-discipline. Bromley, Boran and Myddelton (2007) describe the challenges of assessing these attributes (and recommend a methodology that combines competence modelling with training-needs analysis within a comprehensive self-assessment tool).

We should also, at this stage, touch on one other aspect of higher education assessment and selection. In some circumstances, students are selected to higher education programmes on the basis of an assessment of their existing values, attitudes, dispositions and related aptitudes. The Undergraduate Medicine and Health Sciences Admission Test (UMAT), for example, is designed ‘... to assess general attributes and abilities gained through prior experience and learning; specifically, the acquisition of skills in critical thinking and problem solving, understanding people and abstract non-verbal reasoning. These abilities are considered important to the study and later practice of professions in the health sciences’ (http://umat.acer.edu.au/index.php [accessed January 20, 2009]). Arguably, devices such as UMAT are used when more established, outcome-oriented forms of assessment prove unable to address the forms of learning that higher education institutions most value, or where these traditional forms of assessment are unable to sufficiently categorise applicants.

It would be remiss not to include at this stage some comment on teachers’ own affective stances in these areas. Not all teachers are comfortable with precise descriptions of intended learning outcomes (Hussey & Smith, 2003). Many teachers are cautious about teaching and assessing in the affective domain (Shephard, 2008). Teachers do not all agree on the value of e-learning and use of technologies in teaching contexts and display a range of emotional responses when expected to do so. Salmon (2005) suggests ‘teaching online has almost nothing to do with computers and everything to do with time, motivation, knowledge and the new agency of cyber-experience, as well as good appropriate teaching’ (pp. 214–215). Smyth (2003), discussing broader aspects of change in education, describes the human and contextual factors bound up in the experiences of teachers as they engage in development for change. These are all highly contested areas of practice and enquiry in higher education. While excited about these new research contexts, the author of this paper confides that he is often anxious about his ability to remain as up-to-date with new technologies as many students are.

**Assessment paradigms in higher education and beyond**

This section seeks to establish the broad characteristics of categories of assessment, on either side of graduation, alongside observations about how ICT may support further exploration and development.
Many conventional texts on assessment categorise it as essentially standards based or measurement based. The former attempts to describe the characteristics of individuals in relation to some defined standard and forms the basis of criterion-referenced assessment processes. The latter describes these characteristics so that individuals may be compared with one another and forms the basis of selection processes described as norm-referenced assessment. The literature on these forms of assessment is substantial and, in a higher education context, has been interpreted, for example, by Biggs (1999). More details on each are provided later in the paper, but at this stage, it is important to note that some authors, also with a focus on higher education, identify other forms of assessment that usefully fit other categories. Construct-referenced assessment and ipsative assessment processes are also considered.

It is also important to note that the literature on assessment in higher education and teachers’ understanding and perceptions of assessment are not harmonious entities. Recent research by Taras suggests that ‘lecturers’ understanding of assessment terminology and relationships reflects the fragmented theoretical and practical frameworks available’ (Taras, 2008, p. 172). Moreover, research by Shephard, Warburton, Maier and Warren (2006), with a focus on computer-assisted assessment, suggested that several key aspects of responsibility within this area of assessment were only loosely defined.

**Will I get into a university? Will I get that job?**

Where the number of places is fewer than the number of applicants, processes have developed to allow selectors to identify the most promising applicants. In general, these processes seek to rank the applicants and to identify each individual in relation to all others. Whether the process involves ranking students by their ‘National Certificate of Educational Achievement’ or ‘Advanced’ level scores, or by their performance in a specific test, or by reading their curricula vitae, the purpose of such norm-referenced assessment is generally to determine who is the best suited for the opportunities available. Biggs (1999) provides a thorough analysis of why such assessment processes, often involving normal standardisation transformations and frequently undertaken in higher education, are inappropriate assessment devices within higher education. (Biggs makes the point that assessment in higher education is, or should be, all about assessing change, or learning. Norm-referenced assessment seeks to determine the characteristics of people at some, fixed, point in time.)

Nevertheless, anyone who has endured the task of shortlisting candidates for a popular job will be aware of the need to rank applicants in these circumstances. The use of portfolios for selection and assessment will be addressed in more detail later on, but it is interesting to note that these instruments are used to select candidates for a job. A recent advertisement to recruit a university lecturer required applicants to submit a teaching portfolio (electronic or otherwise was not noted; *The Australian Higher Education*, Wednesday, 6 August 2008). As described previously, UMAT provides a selection-enhancing service at the undergraduate entry stage. e-Technologies support this process, but not in a particularly exciting way.
Did I achieve that learning outcome? Will I pass or fail? What grade will I get for my degree? Will I get promotion? Will I keep my job?

Where the major aim of assessment is to determine if individuals have achieved any particular standard of attainment, generally definable in terms of abilities, then criterion-referenced assessment processes are most frequently used. Student assignments are marked in relation to predetermined criteria. If all the students match the criteria, they all pass. Where criteria are arranged in a hierarchy of levels (often using descriptive assessment criteria), then different individuals can be placed in different categories. Similar philosophies apply to assessment, broadly defined, at other levels. In many professions, higher-education teaching in particular, promotion depends primarily on an individual’s attainment of identified characteristics (e.g., those of a senior lecturer, often described in terms of teaching and research performance). Criterion-referenced assessment has become a central paradigm of academic life in higher education (see, e.g., Rust, 2002, for student assessment). It is also an essential tenet of the change process that has been extant in higher education in recent years, focusing on student-centred teaching. New lecturers in higher education are almost universally urged to adopt constructive alignment by creating a link between an intended learning outcome, a teaching/learning activity and an assessment. Although almost universally adopted, these processes are not universally appreciated. There is a substantial literature on aspects of the use of intended learning outcomes that emphasises circumstances where their use may not be appropriate (see, e.g., an analysis by Hussey & Smith, 2003).

Criterion-referenced assessment also underpins aspects of work-based accountability and judgement in assessments following graduation. In the teaching professions, for example, the use of student perceptions of teaching, to judge the performance of teachers, is widespread and often judgements relate to preconceived criteria of what percentage of students might be expected to ‘value’ a teacher’s teaching (see, e.g., Johnson, 2000). Whether 70% or more students perceive their teacher to be effective as a teacher is an often challenging, and challenged, aspect of professional accountability.

It is likely that it is in this broad area of criterion-referenced assessment that most progress to automate the processes of assessment has been made. Much of this development, however, relates to objective testing for cognitive skills and abilities and need not concern us here.

With an emphasis on the more challenging aspects of affective learning, the key innovations in assessment are in the domain of portfolios. There is an immense and developing literature on portfolios, focusing on their dual and occasionally conflicting roles to support development and judgement. Baume (2007, p. 1) provides a summary of many of the advantages to students, and to teachers, of using portfolios: ‘In producing a portfolio, the student assembles smaller pieces of work into a large whole; makes connections among the items of work they have done; and gives a critical overview of their work and learning. In marking a portfolio, the lecturer sees a coherent and reflective picture of the student’s work and development.’ The creation and maintenance of
a personal portfolio obliges its creator to, for example, analyse and critically reflect on evidence of a wide range of previously agreed criteria, relating to skills, application of knowledge, and if required, particular approaches indicative of values, attitudes and dispositions. Much of this debate and support is duplicated in the e-domain. A survey of e-pdp (professional development planning) and e-portfolio practice in UK higher education has recently been reported and interpreted by Strivens (2007). Strivens, in addition, reports that there is considerable interest in the integration of e-portfolios, virtual learning environments and student record systems within institutions. The broad area of professional development planning does seem to be slowly, but noticeably, progressing in higher education, and e-portfolios, with their substantial integrative affordances, no doubt have been greatly instrumental in this.

Interestingly, much of this debate and development is also duplicated beyond graduation. The professions, and their associated learned societies, have also espoused portfolios for professional development and promotion purposes. The teaching profession is particularly notable in this respect, and it is possible that there are more learned papers on the teaching portfolio than on any other.

While portfolios may provide a vehicle for the assessment of individual criterion-referenced outcomes, arguably, their major and most advantageous use is to assess something broader and akin to ‘the whole package’. This is addressed in the next section.

Do I deserve a PhD? May I join your club? Do I belong?

Some would argue that the PhD examination process is also a form of criterion-referenced assessment. The criteria, however, are often very vague (and may convey no more than the expectation of a novel contribution to knowledge in the discipline). Others (eg, Wiliam, 2008) argue that the assessment seeks to identify the candidate’s attainment of a socially constructed view of what it is to be a PhD. A term that is rapidly gaining credibility for this form of assessment is construct-referenced assessment, and in time this term may come to be used more widely, wherever membership of some exclusive community is in question.

In an educational context, and more broadly within the social sciences, entry to particular communities has become a focus for research within the paradigm of communities of practice. The literature around communities of practice provides rich insights into matters of recruitment, induction and participation in communities (see, eg, Wenger, 1998). Notions of assessment in this domain may be informal, but some communities do develop socially constructed descriptions of practitioners and of practice and seek to use these descriptions to manage the entry of new recruits and to guide their development. Frequently, these descriptions are in the form of abilities or outcomes, and occasionally they are also in the form of ‘values sets’. Tertiary-education teaching provides one example of both forms of description. The UK’s Staff and Educational Development Association (SEDA), for example, offers a professional development framework that identifies core and specialist development outcomes and a common set
of values (SEDA, 2008). Also in the UK, the Higher Education Academy (HEA) promotes its Professional Standards Framework, providing among other things ‘a means of demonstrating to students and other stakeholders the professionalism that staff bring to the support of the student learning experience’ (HEA, 2008, p. 1). Entry to these communities generally requires potential recruits to demonstrate attainment of the indicated objectives and how the indicated values inform the practitioner’s work.

Difficulties associated with notions of belonging are occasionally highlighted. Anthony and Grevholm (2004), for example, describe an interesting discourse on the identities of mathematics teachers, ‘on signposts and guidelines that enable student teachers to effectively map their developing teacher identity against a vision of what it means to be an excellent teacher’ and on a potential disjunction of identities and related images of the mathematics teacher. It is possible that some disciplines and some professions are particularly vulnerable to situations where social constructs themselves are not agreed.

It is perhaps in this area that e-technologies are having most impact. Several terms from the previous paragraphs emphasise the social construction of knowledge (eg, ‘what it means to be an excellent teacher’, ‘a common set of values’ and ‘socially constructed descriptions of practitioners and of practice’). Many authors have commented on the extent to which the Internet and related technologies are empowering a wide range of participants to be involved in social construction of knowledge. Wikipedia provides the most obvious expression of a socially constructed entity, but in some senses, it is but the public face of extensive social networking, online and offline. Blogs and social networking sites may provide the purely online elements of social—construction, but the extent to which they build on communications within more established academic processes of conferences, peer review in publishing and editing has not been determined, nor has the contributory balance between e-mail and telephone conversations, for that matter. Lankshear and Knobel (2003) argue that new economies have developed where resources, information and, nowadays, opportunities for social construction are in oversupply and taken for granted while attempts to gain and maintain attention dominate. Within this expanding and noisy information-environment, new forms of rating, selection and assessment are probably necessary and are being developed. These are, naturally, benefiting from the characteristics of the technology that has created the opportunities. While the acceptance of a journal paper may even now depend on the views of one or two peers, contribution to the wider online environment may be informed by, and even decided by, peer ratings. Lankshear and Knobel (2003) comment that online personal ratings, based on those developed by eBay (the online auction), have become integral features of an increasing range of emerging social practices on the Internet.

It is also possible that ‘traditional’ educators will look to e-technologies to find solutions to some of the intransigent assessment problems. It is widely reported, for example, that students find it difficult to write reflectively, but that these skills are highly respected professional attributes (see, eg, Bush & Bissell, 2008). E-Technologies provide many opportunities for online reflective writing within formal education and in wider profes-

© 2009 The Author. Journal compilation © 2009 Becta.
sional and learning communities. This is an active area of research within the context of professional development, e-learning development and exploration of portfolios (see, eg, Powell, 2006). The link to professional skills, life-long learning and ongoing professional accreditation is also increasingly noticed by professionally related researchers. For example, in dentistry, Kardos, Cook, Butson and Kardos (2008) have been exploring the development of an e-portfolio for lifelong reflective learning and auditable professional certification.

Am I better now than I was before?
Ipsative assessment is not widely and consciously used in higher education, except perhaps in the performance aspects of physical education. The notions of self-assessment, self-reflection, personal–professional development and lifelong learning, however, have much in common with the theoretical underpinnings of ipsative assessment, and arguably, the formal processes of ipsative assessment should be more widely acknowledged in higher education. In some respects, this section is really about reflective practice (Schon, 1983) and self-assessment, as much of professional development hinges on the reflective cycle of planning, doing, recording and reviewing. Again, this is an area where educators hope that portfolios will support better assessment in the future. A substantial development at the Glasgow Dental School, for example, introduced a portfolio to promote student-centred learning and reflection and to fulfil the role of a progress file (Bush & Bissell, 2008). The authors comment on how students find writing personal reflections to be difficult, but the authors consider this skill itself something that the reflective process will support within the portfolio.

Many of the attributes of the portfolio, and e-portfolio, have particular meaning at the level of personal development planning, but other e-related phenomena also have particular relevance here. The role of wikis and related collaboration tools in the development of students’ values and attitudes has not gone unnoticed. Bruns and Humphreys (2005) comment on the nature of Wikipedia’s neutral point of view policy. They suggest that this encourages participants to present all sides of an argument and encourages students to develop the value of objectivity and appreciate the advantages of the collaborative process of negotiating over the content with global peers. Others encourage the use of wikis as teaching and assessment tools for these and related reasons (see, eg, Konieczny, 2007).

Is this assessment or selection? What are the criteria?
Practitioners in higher education will be aware of a number of other processes that have elements of both assessment and selection within them. Three examples will suffice and all are high-stakes processes for the individuals and institutions involved.

The major quality assurance process that directs and controls research publication involves peer review. A researcher’s work rarely automatically enters the public domain. Its presence there, and likely impact on the field of enquiry, depends on the work being selected for inclusion by an editor. The selection process is influenced by the assessments of reviewers, themselves generally experts in the field. Various levels of
anonymity are applied within these processes. The area has been substantially reviewed by Rowland (2002) and by the UK’s Select Committee on Science and Technology (House of Commons) (2004). It is rare for the criteria for acceptance to be detailed by the editor and reviewers generally provide an individual estimate of the value of the work against their own personal constructs. Similar processes apply to the allocation of research funds. Individuals and groups submit applications for funding, generally to contestable but oversubscribed funds. Applications are assessed by peers and selected for funding on the basis of these assessments. Often, the criteria include an assessment by the reviewers of the likely impact of the research on advances in the field of enquiry—a difficult feat by any measure. Peer review is also important in the related process of determining the ranking of individuals and institutions in a research assessment exercise (Research Assessment Exercise in the UK and Performance Based Research Fund in New Zealand), but in some systems, and for some disciplines, it is supported by sophisticated and generally quantitative assessments based on citation indices. The work of individuals, and in some systems, of institutional groups of individuals, is compared on the basis of the extent to which individual elements of research output are cited by others in peer-reviewed publications. Academics may argue about the validity and reliability of these assessments and the fairness of the selection processes that result, but there is no doubt about the extent of, or influence of, these forms of assessment in higher education. It remains to be seen how e-technologies will influence these developments that include aspects of norm, criterion and construct-referenced assessment, but it seems inevitable that a central feature of the e-contribution will relate to new forms of dialogue and analysis that they afford.

Is this assessment or evaluation? What is the difference?
The terms assessment and evaluation are used interchangeably by some in higher education but not by others. For many, assessment relates to the performance of individuals, while evaluation relates to systems or processes. Both terms are widely used outside of education, and education has no exclusive mandate to define the limits of each. Nevertheless, education probably has much to learn from those who assess or evaluate in other fields of endeavour. While education has struggled with its own development of tests, examinations and assignments, the academic field of evaluation has evolved, sometimes in conjunction with education, but sometimes not.

Guba and Lincoln (1989, 2001) provide an informative historical description of four generations of evaluation. The first generation was applied to education and was developed primarily to measure various attributes of school children. The school teacher might be expected to be most interested in the performance of individual children in school tests, but the evaluator was interested in the performance of the test, of the teacher and of the school. The second generation of evaluations focused on providing rich descriptions, and the third focused on the need for evaluators to make judgements about the quality of systems and processes. Guba and Lincoln argue the need for a fourth generation of evaluation, one that emphasises negotiation between all stakeholders in determining the intrinsic value of what is being evaluated and that is built on the broad paradigm of constructivism. Constructivism, of course, itself underpins many
elements of contemporary education and assessment. Constructive alignment and notions of intended learning outcomes are, by and large, constructivist entities. There should, therefore, be much in common between the philosophies of constructivist educators and fourth-generation evaluators, but there are in fact substantial differences in both approach and principle. Fourth-generation evaluators recognise that what is being evaluated has no foundational reality or objective truth and its value depends on the approaches and experiences of those engaged in the evaluative activities (whereas much assessment is based on a more positivist approach, where the notion of a ‘correct’ answer is at least entertained). They also argue that evaluation itself depends on processes of interpretation and logical argument, centred on negotiation that involves all stakeholders. Conventional educational assessments could be managed as fourth-generation evaluations, but they rarely are. Teachers could negotiate with students about the nature of the learning programme and attainment of agreed objectives but this approach is certainly not widespread in higher education. There is, however, an increasing tendency in higher education to adopt self-assessment approaches, often involving learner and teacher as equal partners. Bromley et al (2007) describe an interesting example at the post-graduate level. But it is at the level of external examination (or external assessment, or external moderation) that a clear comparison between educational assessment and evaluation exists. External assessors and their roles vary widely, but they are often most valued for providing oversight on the complete assessment process. They comment on the exam questions, they provide feedback to the teachers and their institution on links between teaching and assessment, they interview the students and the teachers, and they compare processes and outcomes in one institution with those in others. Jackson comments that for the system in the UK, ‘There is a shared belief in UK higher education (HE) that external examining is an important, valued, skilled and useful peer review process. It provides impartial advice to institutions to ensure that students are treated fairly, to enable teaching teams to understand their standards better, and to improve the quality of the education which they provide’ (Jackson, 2004, p. 3). External assessors have, arguably, been working to the basic tenets of fourth-generation evaluation for some time now.

Many of the issues addressed in the previous paragraphs come together in this area. Traditional assessment processes have not proved particularly suitable for recording dispositions, achievements or learning in the affective domain. Approaches new to education, but not necessarily new post-graduation, such as the portfolio, are starting to have impact and are yielding processes where these attributes can at least be recorded. But the processes then have more in common with evaluation than with conventional assessment. In some ways, we have completed a circle. Bloom, Hastings and Madaus (1971), in their early work, identified the difficulties associated with assessing affective outcomes in individuals. They provided suggestions of how courses that promote the acquisition of values, attitudes and behaviours can be evaluated to measure the degree to which they enable students to acquire the required outcomes. They advised that evaluation can be achieved on a group-wide basis and that this approach avoids some of the really difficult issues in assessing values and related affective outcomes. The situations, tools and techniques involve the use of surveys, ques-
tionnaires and semantic differential techniques. We do not need to search far for examples of successful applications of these techniques in online settings as they often use quantitative processes that work well in digital settings. Certainly, e is not only for exploration, but it is also for evaluation.

References


Shephard, K. (2007). Is it the role of higher education to teach students to value our environment, and if it is, how do we do it and how well do we do it? 2nd International Conference on Education for Sustainable Development: Graduates as Global Citizens, Bournemouth, UK.


