

Assessing the Complexity of Professional Achievement

Mantz Yorke



Mantz Yorke's early career was in teaching and teacher education, after which he turned to staff development and educational research at Manchester Polytechnic. He then spent six years as a senior manager at Liverpool Polytechnic followed by two years on secondment as Director of Quality Enhancement at the Higher Education Quality Council. He returned to his institution in 1994, continuing as Professor of Higher Education. Following retirement in 2005, he is Visiting Professor in the Department of Educational Research, Lancaster University.

He has worked on various projects related to graduate employability, the most significant of which involved membership of the Enhancing Student Employability Co-ordination Team [ESECT] whose activities spanned higher education in England (and spread further afield). This work led to his general editorship of the Learning and Employability series of publications by the Higher Education Academy. In addition to his work on employability, he has researched, presented and published on various aspects of higher education bearing on 'the student experience',

Synopsis

Professional achievement is not limited to the well-established professions since professional behaviour is required in a wide range of occupations and voluntary work. This situational variety makes the detailed specification of achievement standards (as some are seeking) extremely – and in most circumstances prohibitively – challenging. Approaches to assessment that are rooted in ideas of scientific measurement (which inflect a lot of assessment in contemporary higher education) are inappropriate to the assessment of professional achievement and need to be replaced by an approach based upon professional judgement. This can be characterised in terms of a shift away from realism and towards relativism in assessment. Two consequences are the need to take a critical perspective regarding the merits of technicalities of assessment that reflect psychometric thinking and, where an overall grade is involved, the need to consider the appropriateness of privileging assessments of academic work over those of professional achievement.

The challenges faced in work and volunteering range from the routine to the novel. The professional has to deal with both, often in situations in which a 'good enough' solution is the best that can be achieved. The assessor has the task of judging the achievement with respect to the prevailing circumstances. Assessors from workplaces often prefer to make broad judgements of achievement. A key issue is the capacity of assessors to make sound judgements. In some areas of higher education assessors are well-versed in assessing professional achievement: in others there is less expertise in such assessing, and hence there is a strong argument for relevant development work.

Assessing professional achievement is demanding on resources. Assessment regimes have to be realistic in terms of what is practicable and what limitations follow regarding what can formally be warranted.

Professional achievement and its assessment: an orientation

The development of professional capabilities is nurtured in many ways – through a variety of curricula in higher education, through experiences and award programmes in parallel with higher education, and through learning in the workplace (both formal and informal). Demands of politicians and employers have for a long time focused on the need for graduates to demonstrate their employability. Though these demands have typically been made with reference to young graduates emerging into the labour market, they have a resonance that extends beyond this to encompass the upskilling of the labour force in general. The USEM approach to employability (Yorke and Knight, 2006 and described later), although focusing on undergraduate education, was nevertheless developed with a concern for graduates' career trajectories.

There have, of course, long existed various programmes in which the workplace and higher education have had a shared interest. Work-related programmes have become increasingly significant internationally as governments have sought to improve their economic position vis-à-vis their competitors. In 2000, the introduction of foundation degrees¹ made explicit the UK government's desire that people in mid-career would develop their capabilities as professionals through a deliberate blending of academic and workplace experiences.

The challenges facing professionals are sometimes routine, and require routine solutions, but often are characterised by unpredictability, such as when a social worker is confronted by hostility at a domestic residence. The professional, faced with unexpected situations, has to call rapidly on a repertoire of possible responses and to make a swift judgement regarding the best line to take. The professional has to deal with the unfamiliar, and at times has to operate outside their comfort zone.

Professional practice is subject to assessment in various ways: it might be formal as in a higher education programme; it might be formal in the workplace (as in a performance appraisal); or it might be informal in the workplace (as when a manager comments on an act of practice). What brings these assessment situations together is the necessity for those making judgements on performance (assessor(s)) to **respond** to what the professional has done in the context in which they are doing it, given a broad frame of expectations, and to come to a judgement about the appropriateness of the professional's behaviour in the prevailing circumstances. This constitutes an approach to assessment that can be characterised as predominantly relativist, in contrast to a predominantly realist approach in which expectations are tightly specified and the assessment focuses on the extent to which the person satisfies explicit criteria. Assessment of professional achievement is, in general, a process in which the available evidence is weighed although, given the sometimes fragmentary nature of the evidence, also contains elements of 'connoisseurship' (Eisner, 1985) may intrude². Whilst neither the relativist nor the realist approaches to assessment exist in a pure form (except in very 'local' circumstances), they constitute different paradigms which have different implications for the technicalities associated with assessment, for the recording and reporting of professional achievement, and for the education of assessors.

This chapter addresses a number of challenges related to the assessment of professionals' achievement.

Variations in settings for the development of professionalism

There is a spectrum of settings in which professionalism can be nurtured via higher education (Figure 1). At one end are programmes that are designed around workplace experiences: these include bespoke company/higher education institution programmes, professional doctorates, and foundation degrees. Many of these involve

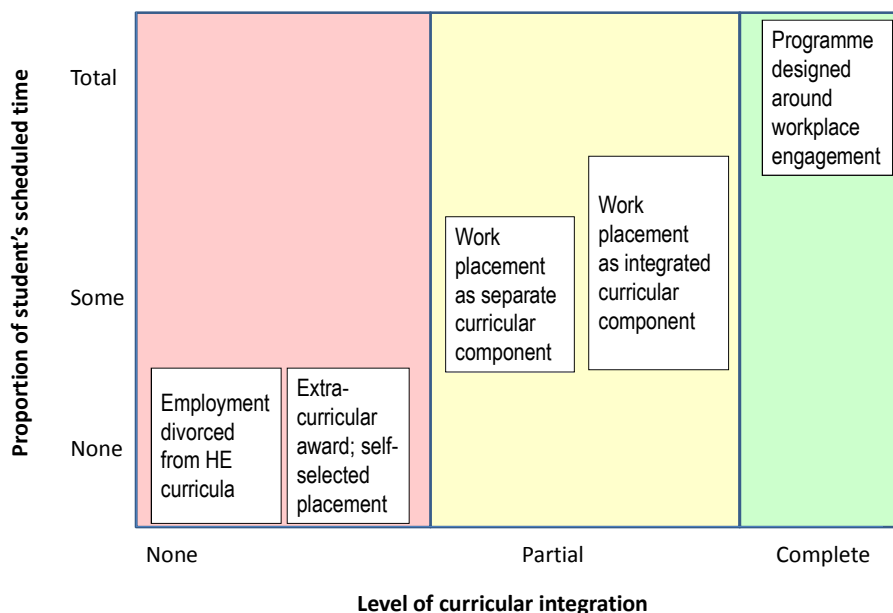
¹ Initially in England and Wales; later in Northern Ireland.

² See Knight (2005) for discussion of differing approaches to assessment.

Assessing the complexity of professional achievement : Learning to be Professional through a Higher Education e-book
<http://learningtobeprofessional.pbworks.com/>

domain-specific knowledge, understanding and skilful practices³. In the middle are work placements designed to be part of curricula: the work placement may be integrated into the curriculum, or it may be a 'stand alone' curriculum component. Willis' (2011) research demonstrates variability of the expected 'skills mix' between departments in a university. At the other end of the spectrum can be found student development awards that recognise and value the development of capabilities, qualities and dispositions relevant to being an effective professional. These awards derive from extra-curricular activities that are not directly related to specific employment possibilities and do not require domain-specific knowledge, understanding and skilful practices (Jackson, 2010a). This context for learner development is probably an under-exploited resource. Yorke (2003) argued that, provided part-time employment could be given a sound academic underpinning, it could be drawn upon for learning in first-cycle higher education programmes. With a slightly different slant, it would seem that part-time employment could similarly be exploited in the development of professional expertise.

Figure 1. A representation of the relationship between workplace experiences and the level of curricular integration.



And of course the opportunities for developing professionalism in the workplace are legion, even if they often do not have a formalism akin to that of programmes in higher education.

The USEM account of employability

Employability, capability and professionalism are overlapping constructs that share an intention that people should develop capacities to act effectively in the world.

³ To the distinction between domain-specific conceptual knowledge ('knowing that') and domain-specific procedural knowledge ('knowing how'), Billett (2009) adds dispositional knowledge ('knowing for'), which moves attention towards the arena of values. *Assessing the complexity of professional achievement: Learning to be Professional through a Higher Education e-book* <http://learningtobeprofessional.pbworks.com/>

The 'USEM' approach to the development of employability (Yorke and Knight, 2006) was itself influenced by four broad and interlocking constructs, and the decision was made to avoid detailed specification since USEM needed to be interpretable and adaptable by the varied subject disciplines in higher education. The constructs are:

- **Understanding**
- **Skilful practices in context** (so labelled in order to acknowledge that academic and work-related achievements are situated in particular contexts, and to avoid the undesirable connotations of 'skills')
- **(self-)Efficacy and personal qualities**
- **Metacognition.**

The development of USEM was influenced by the Higher Education for Capability movement led by John Stephenson from the late 1980s. Stephenson (1992:2) described 'capability' in the following terms:

Capable people have confidence in their ability to (1) take effective and appropriate action, (2) explain what they are about, (3) live and work effectively with others and (4) continue to learn from their experiences both as individuals and in association with others, in a diverse and changing society.

Capability is a necessary part of specialist expertise, not separate from it. Capable people not only know about their specialisms, they also have the confidence to apply their knowledge and skills within varied and changing situations and to continue to develop their specialist knowledge and skills long after they have left formal education.

A difficulty with 'capability' was that, at that time, it lacked theoretical and empirical referents to support its intentions, and hence it struggled to gain leverage beyond a pragmatic coterie in higher education⁴ USEM, in contrast, was able to call on a variety of resources which enabled it to attain a greater degree of acceptance, even in areas where this might not have been expected.

'Employability' tends to be treated as if it relates solely to young people being prepared for the labour market. As is very clear from Stephenson's description of 'capability' (though less prominent in the USEM account of employability), 'employability' is a lifelong construct. However, when people are in workplaces or engaging in voluntary activity, 'employability' is probably better viewed through the lens of professional learning and development.

There is a whiff of individualism in Stephenson's description of capability, despite the fact that the notion of 'working with others' is invoked in his chapter. This may be more an accident of presentation than a matter of design. However, working with others is not followed through in terms of assessment, nor was it discussed in Yorke's (1998) later chapter on the assessment of capability: this is perhaps a side-effect of the tradition in higher education that assessment focuses on individuals' achievements (and which reflects its certification aspect).

Eraut's (2010) sees capability in terms of organisational levels. At the individual level the key is personal knowledge (what persons bring to a situation that enables them to think, interact and perform). One can infer both the complexity of, and the significance of the social dimension in, professional learning when Eraut (2010, p.47) argues that the following all contribute to personal knowledge:

- Codified knowledge in the form(s) in which the person uses it
- Know-how in the form of skills and practices

⁴ The research of Richard Greene (2004) has begun to alter this.

Assessing the complexity of professional achievement : Learning to be Professional through a Higher Education e-book
<http://learningtobeprofessional.pbworks.com/>

- Personal understandings of people and situations
- Accumulated memories of cases and episodic events [...]
- Other aspects of personal expertise, practical wisdom and tacit knowledge
- Self-knowledge, attitudes, values and emotions.

The list (which also acknowledges learning from experience⁵) maps quite well on to Stephenson's notion of capability and to the USEM approach to employability. Where Eraut goes beyond these formulations is in his concern to construe capability in terms of how, at the levels of the team and the organisation as a whole, the respective collectivities address the issues that face them.

The 'messiness' of workplace problems

With considerable caution, because subject disciplines vary so widely, it can be argued that problems set in higher education and those in professional life tend to differ in some – occasionally all – of the ways summarised in Table 1. The danger is that, because the differences are set out baldly in the Table, the necessary tentativeness inherent in the Table might be overlooked. The point of the Table is to suggest some of the ways in which differences might arise – it is for readers to determine the extent to which the differences apply to the circumstances of interest to them.

Table 1 Some contrasting features often found in problems set in higher education and in those faced in professional life. Note: the first column is based on Hedlund and Sternberg (2000, p.137).

Problems set in higher education are quite often characterised by	Problems faced in professional life are quite often characterised by
<ul style="list-style-type: none"> • being deliberately formulated • being well defined • having a 'right answer' ... • ... and a preferred way of reaching it • the availability of most, if not all, of the relevant information • being of limited intrinsic interest • their detachment from ordinary experience • adequate time • being posed to the individual 	<ul style="list-style-type: none"> • 'happenstance' • 'messiness' • multidisciplinary • incompleteness of information • the pragmatic need to satisfice – i.e. to reach a 'good enough' solution within the time and resources available, rather than a perfect solution • requiring the involvement of more than the individual

Professional behaviour implies, at times, tackling the unfamiliar

Stephenson (1992:4ff) notes that much of the educative process is focused on dealing with familiar problems in familiar contexts. Relatively standard solutions can be applied. This is represented in the shaded call of Table 2, which is a tabulation of Stephenson's graphical depiction.

⁵ The late Norman Evans was a strong advocate of learning from experience, through his work for the Learning from Experience Trust. Learning from experience is not necessarily 'good', since the experience may constitute poor professional practice. *Assessing the complexity of professional achievement : Learning to be Professional through a Higher Education e-book* <http://learningtobeprofessional.pbworks.com/>

Table 2. The context – problems matrix adapted from a Figure by Stephenson (1992)

	Familiar problem	Unfamiliar problem
Familiar context	Standard responses are likely to apply, though the respondent must take care not to be complacent about the assumption of familiarity.	Self-evidently, this requires the respondent to move beyond the comfort of existing capabilities, and to find additional support.
Unfamiliar context	A solution known to apply in a familiar context may not transfer to an unfamiliar context.	The most challenging situation for the respondent. Stephenson refers to 'independent capability' when working on problems in this cell ⁶ .

The shaded cell also tends to signal the respondent's dependency on the authority of others: dependency can be seen as one pole of a bipolar construct running between acquiescence (to authority) and autonomy. Various theorists, although using different terminology and focusing on younger students, point to a developmental trajectory running from acquiescence to autonomy, whose directionality should be facilitated by higher education (see Kohlberg 1964, King and Kitchener 1994, Perry 1998, and Baxter Magolda 2009).

A connection can be made with the stages in professional development proposed by Dreyfus and Dreyfus (2005). The novice begins their developmental trajectory by following rules that have been laid down (acquiescence); learns to appreciate the significance of context in the application of rules; learns to cope with the sheer complexity of many practical situations; and ends up as a proficient or expert professional who is able to select, from the plethora of information, an appropriate professional response (autonomy). Gerard Manley Hopkins, delighting in the flight of a kestrel in his poem *The Windhover*, wrote of 'the achieve of, the mastery of the thing'. The true professional, it can be argued, commits to seeking the highest levels of mastery and achievement, and is not content with acknowledging true expertise whilst maintaining merely a humdrum level of performance. It is the person's determination, perhaps, that is the key to becoming an expert professional.

The Dreyfus brothers' account emphasises cognition⁷ and, as Eraut (2009) remarks, it is individualistic and conservative since it pays scant attention to the involvement of the professional in various social groupings. Greene (2004), whilst acknowledging the importance of the cognitive in eight general capabilities of highly effective people, gives a much stronger emphasis to personal qualities and attributes in personal success⁸. Not only are highly successful people very self-aware (i.e. they act very metacognitively), they also have the capacity to project their capabilities in such a way as to influence others.

As soon as unfamiliarity in either context or problem raises its head, standard responses may be inadequate. A familiar problem but in a different context may require a different response: to take a relatively dramatic example, what might work in a European context may well not transfer to an African village – cultural and material conditions may simply be too different. An unfamiliar problem in one's 'home' context would clearly involve the need to search for support outwith one's current capabilities. Once one moves beyond the comfort zone of familiar problems requiring familiar solutions, there is an element of risk as regards the performance. Risk has to be managed during the learning experience (for example, no-one would advocate placing hospital patients at the mercy of a tyro). However there is another dimension to risk – that the person may stumble badly when attempting something at the edge of their capability, but may gain from the experience. From an assessment

⁶ See Cairns and Stephenson (2009, p.13).

⁷ Despite Eraut's (2009, p.2) remark that the Dreyfus brothers' account was developed as 'an antidote to the hyper-cognitive perspective on learning developed by cognitive scientists in the late 1970s and early 1980s'.

⁸ Sternberg (1997) had previously argued forcefully that personal qualities make a significant contribution to achievement.

Assessing the complexity of professional achievement : Learning to be Professional through a Higher Education e-book
<http://learningtobeprofessional.pbworks.com/>

perspective there is a dilemma: put bluntly, can learning from failure count as success? Jackson (2010:25) makes clear his position on the issue:

[The] curriculum must ... prepare [students] for and give them experiences of venturing into uncertain and unfamiliar situations and enable them to encounter and learn to deal with situations that do not always result in success and which require resilience and persistence to overcome difficulties and meet the challenge. It should not penalise 'mistakes' or see the lack of an immediately successful outcome as a failure to achieve but see the process of trying to achieve in such uncertain and difficult territory as a successful achievement in terms of learning for the future.

Assessing professional behaviour

Assessing professional behaviour involves at least three kinds of achievement:

- what knowledge, understanding, values and capabilities the person can demonstrate when engaged in their professional work;
- what s/he achieves in work situations; and
- how s/he goes about her/his work.

The blending of these kinds of achievement indicates – as Eraut (2010) is at pains to point out – how complex professional behaviour is.

There is a loose fit between the three kinds of achievement and the USEM account of employability. It needs to be appreciated that the USEM account was developed with broader intent than dealing with 'employability'. USEM was developed as a response to political demands on higher education to develop employability in students but the aim was actually wider – to encourage students to become what might be termed 'effective operators in the world' in respect of workplaces, the community and personal networks of various kinds, as well as developing 'the whole person'⁹.

Knowledge and understanding fit the U of USEM quite closely, but have to be appreciated as relating not only to the learning of academic content but also to the learning of how organisations work in practice. The latter does, of course, have academic implications as well as socio-cultural ones. The academic learning in the subject discipline might also encompass the development of skilful practices within the discipline, such as conducting infra-red spectrometry to a high standard – part of the S of USEM.

Performance in work situations relates to the S of USEM in that it is concerned with the skilfulness of practice. A social worker, for example, needs not only to be able to apply in practice what has been learned academically¹⁰ but also to attune the approach to application to the person(s) with whom they are dealing – i.e. to demonstrate 'emotional intelligence' (Salovey and Mayer, 1990; Goleman, 1996). The skilfulness of practice subsumes academic and social dimensions.

How the person goes about their work is intimately – and inextricably – linked to their practice. The 'how' aspect of the work involves both a slew of personal qualities and attributes (the E of USEM) and – perhaps as part of E but given separate attention in USEM – the capacity to operate metacognitively (the M), in which reflection in, on and for practice is of key significance.

⁹ The balance between the various components would be expected to vary according to the background of the student. For example, the developmental needs of a mature student with a wealth of life-experience are rather different from those of a school-leaver aged 18.

¹⁰ 'Academically' is not limited to learning from an academic institution: it can occur during practical engagement or through self-determined study.

Highest or normal level of performance?

In assessing a person's level of achievement, the context of the assessment is a significant factor. The person may perform in a different way (better or worse) if they know they are under a high-stakes test from when they are working 'naturalistically' on a day-to-day basis (see, for example, Hays et al, 2002; Schuwirth et al, 2002). Performance under formal test may not necessarily be a good indicator of how the person will perform under normal working conditions. Under 'workaday' conditions, the person will in many instances be under observation for assessment, whether the assessment is being undertaken with formative and/or summative intent. As part of a discussion of 'capability' that extends beyond Stephenson's (1992) conception, Eraut (2010:48) construes 'capability' in terms of an inference from a series of performances, each of which is context-dependent. In so doing, Eraut is pointing towards a technical issue, discussed later, relating to generalisability.

Approaches to summative assessment

Jessup's (1991) model for the assessment of National Vocation Qualifications [NVQs] in the UK (which owed a lot to the instructional objectives approach adopted by Mager (1962) and others) proved inadequate to deal with the complexity of performance where contextual matters had to be taken into account. Whereas performances could be specified in detail for basic vocational achievements such as wiring of electrical equipment, higher levels in the NVQ framework could not be specified with equivalent precision because of the 'situatedness' of performance. The higher levels of NVQ, for example, involved elements of managing others – with the obvious intrusion of variables (especially people) that could not be treated as independent of the context in which the performance was happening.

In higher education over the decades, there has, in curriculum design, been a shift in focus from educational objectives to intended learning outcomes, reflecting a greater emphasis on learning. Both, however, tend to be strongly teacher-determined. For the moment, any reservations about teacher-determinism are set aside, and the rather old-fashioned terminology of objectives is used for convenience in presenting a typology of objectives/learning outcomes and a commentary on assessment (Table 3).

Table 3. The mode of assessment associated with different kinds of educational objective.

Type of objective	Problem	Solution	Assessment
Instructional	Specified	Specified	Prescriptive
Problem-solving 1	Specified	Open	Responsive
Problem-solving 2 ('Expressive')	Specified/Open	Open	Responsive
Social	Open	Open	Responsive

Instructional objectives are addressed through tight control of specification, expected response (solution) and as a consequence the assessment is prescriptive. The work of Mager (1962) and Jessup (1991), amongst others, is testimony to this line of educational thinking. Problem-solving objectives have two basic forms, depending on whether the problem is specified by the teacher or determined by the student. Both forms imply an openness to a variety of responses – indeed, if the specified problem is accompanied by a particular expected solution, then it is properly characterised as 'puzzle-solving' which is a variant of the instructional objectives approach. If the problem is specified by the teacher, the responses may in some circumstances be constrained, limiting the extent to which there is freedom on the part of the student's response. Eisner (1985) advocated a version of Problem-solving 2 in the field of Art & Design, arguing for 'expressive objectives' which were much more general than those espoused by Mager and allowed students to respond as they saw fit to the challenges facing them. Eisner wrote:

An expressive objective describes an educational encounter: it identifies a situation in which children are to work, a problem with which they are to cope, a task in which they are to engage; but it does not specify what from that encounter, situation, problem, or task they are to learn. An expressive objective provides both the teacher and the student with an invitation to explore, defer, or focus on issues that are of peculiar interest or import to the inquirer. An expressive objective is evocative rather than prescriptive. (Eisner, 1985: 54-5.)

It is clear that the educational encounters set up for the students (with varying levels of specification) were expected to lead to outcomes that could be recognised and assessed *post hoc*.

In discussing Eisner's argument, Popham admitted to being confused by Eisner's use of the word 'objectives', and argued that he [Eisner] was describing learning activities rather than objectives¹¹. It is clear, however, from Eisner's writing that he had in mind broad objectives that offered students freedom to determine, as appropriate to the general framing of the task, their own response from a relatively unconstrained range of possibilities, in contrast to the tight specification associated with a 'hard' behavioural objectives approach.

The disputation over 'expressive objectives' has echoes in current curricular debate. If intended learning outcomes are stated fairly tightly, then the assessment can be close-coupled (and, as a collateral consequence, evidence is readily available to satisfy requirements regarding quality assurance). Professional behaviour, however, cannot be specified tightly: the Eisner argument about providing general intended outcomes against which responses can vary and be subjected to *post hoc* assessment is a closer fit with what is needed.

'Social' objectives are included in Table 3 because, like 'expressive objectives', their achievement emerges as a consequence of engagement with others in a workplace or an academic context. The engagement might be as mundane as working in a fast-food kitchen or as sophisticated as working in a design team for a new industrial bio-reactor.

The key point about the last three rows of Table 3 is that the pre-specification of preferred responses is inappropriate where students have freedom to respond in their own way to the challenges that confront them, and hence that assessment of achievement has to be *responsive* to performances that may be significantly influenced by situational contingencies. These contingencies include, for example, the need to strike a balance between competing demands, the specific requirements of professional and statutory bodies and constraints such as the need to consider the economic implications of potential courses of action.

Yorke (1998) argued that, where an individual's demonstration of practical capability was to be assessed, there was a need to switch from specifying intended learning outcomes in detail (and testing accordingly) to stating the outcomes in more general terms and using assessment as a vehicle for *judging* the extent to which the performance(s) satisfied the general expectations.

The generic assessment question is ... switched from something like 'Can you do this?' to 'How have you satisfied, through your work, the aims stated for your particular programme of study?' (Yorke 1998:181.)

Challenges which can be subsumed under the question 'Can you do this?' imply (to a first approximation, at least) that there is an objectivity to the assessment that locates it within an approach to assessment that is fundamentally realist in character. The 'this' of the question implicitly has a correct (or at least a preferred) response, boxing in the student seeking a grade indicative of success. Challenges in the lifeworld, however, are often complex and involve striking a balance between pressures that are competing and possibly oppositional.

¹¹ See Eisner (1985:60). Eisner did not avoid muddying the distinction between objectives and outcomes when stating in an Epilogue to discussion of his argument:

... an expressive objective is the *outcome* of an encounter or learning activity which is planned to provide the student with an opportunity to personalize learning. It is precisely because of the richness of these encounters or activities and the unique character of the outcome that the expressive objective becomes so difficult to describe in advance. (Eisner, 1985: 69, original emphasis.)

Assessing the complexity of professional achievement : Learning to be Professional through a Higher Education e-book
<http://learningtobeprofessional.pbworks.com/>

Such challenges may not have solutions that can be pulled off the shelf: responses have to aim for optimality in the prevailing circumstances, and with the challenged being time-poor (as often is the case), 'satisficing' (Simon 1957) may be the best that can be achieved. The assessor is then faced with making a *post hoc* judgement as to whether the response is 'good enough' in the situation¹². This represents a more relativist approach to assessment which may nevertheless contain elements of a realist approach.

Lester and Costley (2010) give responsiveness in assessment a different slant when they remark, in relation to work-based learning but with broader import:

... there is a need to assess work-based learning through methods that are adequate, valid, and avoid undermining the nature of the learning, given that it will typically be issue-based, driven by the learner, and transdisciplinary. The aim of assessment is generally to assess learners' progress as 'map-makers' or self-managing practitioners ... not to confirm their conformance as 'mapreaders'. (Lester and Costley 2010: 566)

Assessment before, during and after professional activity¹³

The validity of assessment methods is necessarily tied intimately to the context of the professional activity.

There are three broad contexts that need to be considered:

- when the student is learning the basics of becoming professional in their approach to employment and other situations, but has yet to experience them 'live' as part of the study programme;
- when the person is acting in a professionally relevant environment; and
- when the person is reflecting on their professional activity.

Before engaging in professional activity

Prior to engaging in professional activity, the student can be faced with simulations, case studies and – of course – learning from theory and practice as presented in texts. Assessment methods can include

- behaviour in, and analytical reflections on, simulations
- analyses of case studies
- assignments and examinations of various kinds in which knowledge and understanding (at least) are tested.

During professional activity

During professional activity, much will hang on observation by a professional with sufficient capability in both their professional role and in judging the performance of others. Such methods can include

- Direct observation of individuals (which may involve checklists, as in some medical protocols)
- Direct observation of group activities (where the focus might be on individuals, the group as a whole, or both)

Following professional activity

In the assessment process, activity as a professional can be followed up in a variety of ways. Some of these ways depend on materials produced during, or closely after, the period of engagement. Assessment methods can include

- reflective analyses or commentaries of various kinds, such as logs or essays
- reports on the professional activity (e.g. work placement)

¹² Note that a judgement of 'good enough' does not necessarily imply that it is grudging. A 'good enough' achievement in particular prevailing circumstances could be rated as a 'very good' or even an 'excellent' achievement of the intended learning outcomes.

¹³ This section emerged through discussion with the editor

- portfolios of experience (though these require consideration to be given to both the ethics of an assessor reading what could be a highly personal document, and to the implications for the assessor's commitment of time). An analytical digest of the portfolio content might be more appropriate and practicable.
- presentations relating to the professional activity (e.g. orally, via some form of media, or via posters)
- interviews
- evaluation of artefacts produced during the professional activity.

Given the uniqueness of each episode of professional activity, and the dangers associated with over-inferring from limited samples of behaviour, there needs to be an overarching perspective on summative assessment. The notion of lifelong and life-wide learning trajectories (Eraut 2010:62ff) can be cashed, in assessment terms, by treating each instance of professional action as a window on an episode of practice. Willis et al (DATE TBA) have undertaken both qualitative and quantitative studies which support the 'trajectories' approach.

Academics are, on the whole, best placed to deal with assessments of learning prior to, and following, professional activity. The assessment of actual professional practice is, in some subject areas (e.g. Nursing, Teacher Education and Social Work), shared between academics who have substantial professional experience and professionals in the field. In other subject areas, it is the professional in the field who is best placed to assess.

Realism and relativism in summative assessment

The realist and relativist approaches to assessment can be set up as contrasting 'ideal types' (Table 4). It is unlikely that either can be found in a pure form: each is inflected by the other.

Table 4. Assessment, as viewed from realist and relativist standpoints. From Yorke (2011)

Realist	Relativist	Comment
Standards are objectively defined	Standards are normative and consensual	The realist's objectivity reflects underlying values
Performances can be measured against these standards	Performances are assessed with reference to these standards	The difference here is between measurement and judgement
The assessor is objective and detached	The assessor interprets the extent to which the performance relates to the standards	The underlying distinction is between context-free assessment (the realist's position) and context-relevant assessment (the relativist's position)
Values play no part in the assessment	Value positions are embedded in the norming of standards	
The situation of the student is not taken into account	The context of the assessment is taken into account	
Explicit criteria and rubrics are invoked	There are broad statements of expectations	The issue is the degree of detail that can be encompassed by criteria
Measurements are taken as true and reliable representations of achievement	Assessments are judgements of the extent to which achievements fit with expectations	The distinction is between measurement and judgement
Tasks are set by assessors	Tasks may be selected by students to suit their strengths and interests	The selection of project topics is a feature of both the sciences and the social sciences

In his poem *Among School Children* W.B. Yeats asked: 'How can we know the dancer from the dance?' Much assessment in higher education is focused on the performance essentially detached from the person producing

it. Written productions are often deliberately anonymised in the interests of minimising the chances of bias in assessment. Some productions, of course, cannot be anonymised because the capacity to practise necessarily involves an individual's performance – competence in teaching and social work constitute two examples¹⁵. The chances of bias can be reduced if the role of assessor is separated from those of teacher and supervisor, as is the case in the Foundation Degree in Breast Imaging that is run jointly by St George's, University of London, and Kingston University (see Woolf and Yorke, 2010:16).

The Yeats quotation neatly encapsulates the challenge for the assessor. Whilst some aspects of performance can be separated out and subjected to scrutiny (much achievement in the academic milieu falls into this category), workplace performances are less susceptible to such separation. 'Workplace performance' subsumes a spectrum of achievements ranging from those whose standards can readily be judged against professional specifications (the ability to perform an injection correctly is an example), and those that have to be assessed in a more holistic manner. The former involves the stipulation of tightly-drawn criteria for assessment (no-one would want a medical practitioner who could not inject correctly), but performance as a professional is inherently a more rounded matter. Hence much assessment of professional achievement necessarily lies towards the relativist end of the realist/relativist continuum.

Assessment in the workplace

Thus far, the discussion of assessment has emphasised summative assessment in respect of academic credentials. Assessment in the workplace reflects norms and expectations regarding performance which, in some kinds of organisation, are inflected by those of professional bodies. It has both formal and informal dimensions. The formal is represented in organisational appraisal systems, and has both summative and formative components. Eraut (2009:1-2) remarks, in a paper addressing emergent professionals:

Newcomers often have to learn "How we do things here" without being given any specific objectives or advice. Thus a learning goal might be described by a vague phrase like "being able to do what X does". Even when more detailed advice is given, your learning will still be evaluated by the extent to which you can do what X does, rather than by some indirect and less authentic type of assessment. You may be given sets of objectives or competencies, but the 'real' assessment will be whether your performance meets the expectations of significant others in your workplace.

Although the workplace appears to be primarily concerned with your capability (what you do and how you perform), it is equally important to be able to do the right thing at the right time. In practice this means that you have (1) to understand both the general context and the specific situation you are expected to deal with, (2) to decide what needs to be done by yourself and possibly also by others, and (3) to implement what you have decided, individually or as a group, through performing a series of actions. All three of these processes contribute to your perceived competence. Even if other people are making the decisions, you may still have to interpret their meaning in order to know precisely what is required.

Much assessment in the workplace is informal, even though it may be filtered into the appraisal system. Sometimes this is overtly and deliberately formative (such as when the assessee is judged to have fallen short in some respect), but at other times it may be necessary for the assessee to detect, interpret and act on signals from a work colleague, such as the raising of an eyebrow or unusual gruffness and the avoidance of eye-contact.

Some technical considerations

¹⁵ The exemplification can be extended into other fields. Artefacts from programmes in Art & Design are often related to the student's 'journey' as revealed in a portfolio, rather than assessed in a decontextualised way.

Assessing the complexity of professional achievement : Learning to be Professional through a Higher Education e-book
<http://learningtobeprofessional.pbworks.com/>

Any discussion of summative assessment cannot avoid the challenge to demonstrate robustness. The psychometric, measurement-oriented approach to assessment which is coherent with the behavioural objectives approach to education emphasizes validity, reliability, objectivity and generalisability. The first three of these constructs align with the realist position outlined earlier, and generalisability can be seen in terms roughly analogous to those applying to inferential statistics – if the sampling of behaviours is appropriate, then an inference beyond the actual data (i.e. the generalisability) can be ascribed an appropriate level of confidence. ‘Up to a point, Lord Copper’ as the journalist Henry Boot remarked in Evelyn Waugh’s novel *Scoop*. As one moves from straightforward technical ability (one is, or is not yet, competent in performing a task) to situations calling for integration of capabilities, assessment becomes more a matter of judgement, and hence there is a shift in assessment from realism towards relativism. Guba and Lincoln (1989), in writing about evaluation methodology (rather than about assessment), suggested a different quartet of constructs to replace the psychometricians’ quartet – credibility, dependability, confirmability and transferability. The constructs suggested by Guba and Lincoln offer prompts towards a rethinking of assessment practice regarding professional achievement.

Validity and credibility could be argued to be dealing with the same issue – does the assessment activity adequately reflect what is expected in the way of the assessee’s performance?

Reliability, in psychometric terms, has to be high or else it compromises validity. Shifting towards relativism implies the need to look for an alternative construct that is capable of catering for the variation in assessment circumstances inherent in professional behaviour. Asking whether the assessment judgement is dependable (and requiring evidence on the point) seems preferable. It also asks implicitly about the transferability of professional behaviour from one context to another.

Confirmability, being substituted for a detached objectivity, is a requirement that what is manifested by way of performance, and its interpretation, are not coloured by the assessor’s prejudices: the case for confirmability has to be coherent and consistent, and available to external scrutiny.

For the psychometrician, the ‘reactivity’ of an assessment should be low – i.e. it should not significantly affect the assessee’s behaviour. The link with objectivity is palpable. However, this implicitly sees assessment as a ‘one-off’. In reality, learners often have the opportunity to redeem weak or failing performances until they have achieved a suitable standard (the UK driving test is a commonplace example). The assessee learns from the experience of assessment, and performs accordingly. Much assessment is deliberately reactive in nature. Formative assessment is obviously so. But the summative assessment of professional behaviour is often reactive, too, with the assessee’s performance being appraised and feedback being provided as to how improvement might be made or how they might move on to a higher level of expected performance and undergo assessment to the higher standard. Seen in this light, reactivity is an essential part of the trajectory of development of professional capabilities. Summative assessment is not only about standards reached (judged either against specification or *post hoc* with reference to professional norms), but is also undertaken with reference to the learning trajectory.

One of the challenges faced by higher education is dealing adequately with the limitations that exist in terms of resources and time. Assessors can, with acceptable robustness¹⁶, certify some kinds of achievement, such as academic learning, through the assessment of written productions.

Other aspects of summative assessment are relatively uncontentious, and are merely listed here:

- it needs to be fair (though, as Stowell, 2004, pointed out, fairness is not unproblematic)
- it should not be vulnerable to cheating

¹⁶ Though there is probably a measure of complacency here – see Yorke (2008) for argument of the case. *Assessing the complexity of professional achievement: Learning to be Professional through a Higher Education e-book* <http://learningtobeprofessional.pbworks.com/>

- it should have practical utility and hence its outcomes should be intelligible to those who draw upon them.

The contentious issue of standards

There have been, over the years, various attempts to codify expectations for learning and assessment: Jessup (1991) and Mayer (1992) provide examples from the UK and Australia, respectively. Such codifications have their uses. They help interested parties to clarify what achievements are being sought, and hence facilitate informed discourse. The Australian Universities Quality Agency [AUQA] makes a fair point in its discussion paper of 2009 when it states:

Introducing standardised and understood methods of assessing and grading these [generic] attributes, at the level of difficulty appropriate to the stage of the learning process, ensures that students better understand why they must learn particular things and also provides meaningful evidence to use as part of their future career activities. (AUQA, 2009:12)

However, the problem with such codifications is that they can, all too easily, lead to a 'box-ticking' approach to assessment in which relatively small aspects of performance are totted up to produce an overall assessment. The whole is implicitly assumed to be the sum of the parts. The rule-following novice, in the Dreyfus brothers' categorisation, would probably be well served by this approach.

The Organisation for Economic Co-operation and Development [OECD] has established a pilot project (Assessment of Higher Education Learning outcomes – AHELO) to assess students' 'generic skills' in areas such as critical thinking; analytical reasoning; problem-solving and written communication¹⁷. This is a narrow perspective on 'generic skills', which leaves out of consideration a lot of aspects of professional behaviour. Like the OECD's Programme for International Student Assessment [PISA], which compares countries' overall performance at school level, AHELO is being established with a broadly similar remit at higher education level. Whilst the 'generic skills' being tested in AHELO have an 'academic' relevance, they are also relevant to professional behaviour. They may be necessary for professional achievement, but are manifestly insufficient. Given the limited range of 'generic skills' being tested, it is difficult to avoid the conclusion that the project is driven by what is (roughly speaking) measurable.

Once one gets beyond the status of the Dreyfus brothers' novice, professional achievement demands not only capability in respect of the individual components but also the capacity to integrate them and to draw in knowledge and expertise from elsewhere: the whole is definitely more than the sum of the parts. In an interesting small-scale study that seems to bear out the general 'message' of the Dreyfus brothers' account of professional development, junior medical personnel scored higher than their more experienced colleagues on checklist relating to an objective clinical structured examination [OSCE] (Hodges *et al*, 1999). Presumably they were following 'the rules' closely whereas the more experienced were able to take short cuts in diagnosis. When global ratings were used with the same personnel, the more-experienced outscored the less-experienced.

Eraut (2004:804) pointed to the importance of integration in professional behaviour:

... treating [required competences] as separate bundles of knowledge and skills for assessment purposes fails to recognize that complex professional actions require more than several different areas of knowledge and skills. They all have to be integrated together in larger, more complex chunks of behaviour.

¹⁷ See www.oecd.org/document/22/0,3746,en_2649_35961291_40624662_1_1_1_1,00.html

This chunking of behaviour and the 'situatedness' of professional achievement constitute a significant challenge to those who might wish to codify standards. Political pressure on higher education to demonstrate accountability (evident in the UK, Australia and elsewhere) has instigated action to find ways in which standards outside the area of academic work can be codified and used in benchmarking processes. AUQA wrote, as a lead-in to the quotation cited above:

Most Australian institutions have developed general learning outcomes ('generic attributes'), in addition to the discipline-specific ones which are used to define students' experiences, and therefore achievement, at that institution. These attributes need to be measured and graded. (AUQA, 2009:12)

This falls straight into the trap of fragmentation pointed out by Eraut (2004); it implicitly assumes some form of standardisation across contexts; and it assumes that (*contra* Knight and Page 2007) that a measurement model of assessment is appropriate.

Standards of professional behaviour vary from the specific (e.g. in respect of medical practices, where competence is mandatory) to the general. It is the general that poses a challenge to quality assurance, since it is not amenable to being tied down to narrow descriptors and often the general is applied in contexts where there are no right or wrong answers or ways of acting. The challenge of the general can sensibly be met only by stating the expectations for professional behaviour in broad terms, and judging whether the achievement is an adequate (or perhaps a very good, or an excellent) response to the expectations. The judgement has to be evidence-based, which ought to be sufficient for the purposes of quality assurance. Comparability of achieved standards can only be determined at the broadest level: to hunt for precision in this respect is to chase a chimera.

No-one is an island

Many awards that are made to individuals reflect the contributions of others: how often is it said (sometimes tearfully, as at the Oscars) that the individual sees the award as being made not only to them but also to the team of people with whom they have worked. Sometimes, a duo is more than the sum of the parts. Examples from the UK include the comedy pairings of Eric Morecambe & Ernie Wise and Tony Hancock & Sid James, and the football managerial pairing of Brian Clough & Peter Taylor. The same can be said of teams: readers of Belbin's (1981) book will recall that successful teams are composed of individuals with complementary strengths, and that a team composed entirely of highly creative individuals performed less well than one containing a mix of creative and down-to-earth practical people.

Occasionally, as in the *Times Higher Education* awards, the team is honoured – but the competitiveness inherent in the process leaves runners-up who may well have made significant achievements with no more than a certificate. The 'Change Academy' run by the Leadership Foundation for Higher Education [LFHE] involves cross-sectional teams from institutions in working together on institutional-relevant problems. Whilst the approach has been successful in that institutions have benefited, there is no formal recognition by the LFHE of the teams' achievements. Some teams have, however, been able to draw on their achievements to gain support for further work, which is an informal recognition.

Two questions are prompted:

1. How might an assessor 'pick out' the contribution of an assessee to a group's achievement?
2. Does professional achievement in a group need some form of group recognition?

The first question is probably not best answered by envisaging a single assessor. A number of interested parties have the potential to contribute, including the assessee. One approach might be to require the assessee to construct a case study of their achievements, and for others associated with the assessee's work to provide evaluative commentaries on it (these commentaries could be undertaken contemporaneously or in response to the case study). A non-involved person could then review the evidence and come to a judgement as to the

assessee's achievement. This would be unlikely to be a fine grading – a coarse grade, perhaps a simple pass/fail, may all that can sensibly be awarded. Note that this approach is time-consuming compared with traditional assessment methods but, if the achievement is of sufficient importance, resourcing has to be made available to accommodate the assessment procedure.

The second question takes assessment into a completely different realm. It is one that is well worth asking, but to consider group awards implies some serious thinking about the curriculum in which it is embedded – a task that is beyond the present chapter.

Some practical challenges

The employment literature is full of references to desirable qualities in the workforce like flexibility and adaptability. These kinds of construct are a poor fit with modes of assessment that are dominated by psychometric thinking, since successful behaviour takes the assessee outside the 'box' of pre-specification. Dependability, coupled with transferability (as used by Guba and Lincoln 1989), provides pointers to the assessee's adaptability and flexibility, and the evidence is enhanced where multiple assessments are made, and from different perspectives (this calls to mind the so-called 360° approach to appraisal). It is incumbent on the assessing organisation to integrate the evidence of achievement, and not to leave the various fragments like the un-connected pieces of a jigsaw. Note, though, that the evidence is of past performances and (as the financial services sector is at pains to point out) offers no guarantee of future performance. The evidence may, however, offer a qualitative probability that the assessee may (or may not) be successful in undertaking a new role or task.

A real constraint that applies particularly to higher education institutions and smaller enterprises, is the resourcing of assessment. On higher education courses, there is insufficient time, and smaller enterprises simply do not have the time or expertise to undertake a spectrum of assessments. Ideally, assessment demands multiple occasions, a range of assessors, and a mechanism for bringing all the evidence together and weighing it in the round. An assessment centre probably comes the closest to achieving this, but at the expense of some distancing from 'real-life' performances.

Perceptions of quality assurance tend to push summative assessment towards tight criteria whose application can if necessary be demonstrated to quality auditors. The trouble is that the complexity of professional life (or, on a smaller scale, work-based learning) does not fit this kind of approach. The actual requirements of quality assurance (rather than perceptions of the requirements, which often turn out to be self-imposed constraints – 'quality assurance wouldn't allow it') ought to be more receptive to assessment relying on dependability (as outlined above) rather than on some variant of psychometric reliability.

A major challenge in the assessment of professional achievement is to ensure that assessors have adequate professional expertise in both the field within which the behaviour is being manifested (e.g. law) *and* in educational assessment (note that 'adequate' might be different for the two components). Some questions are:

- How many academics can truly claim to have recent and relevant experience within their compendia of professional expertise, when they come to assess a professional's performance?
- Whereas they possess disciplinary expertise, do they also have the capability to assess, say, the 'soft skills' valued by workplaces?
- Do academics see enough of an assessee's performance to reach a valid judgement?
- Do workplace assessors have sufficient expertise as regards assessment?
- How should assessments from the workplace and academic aspects of programmes be combined?

Academics assessing professional achievement often – and for understandable reasons – have to make do with a limited sample of the assessee's performances. They simply do not have the time to do otherwise. Some have had relatively little recent immersive exposure to workplace environments. Then there is the issue of the extent to which academics have expertise in the assessment of 'soft skills'. For some, assessing 'soft skills' will

be akin to assessing in a foreign language in which they have perhaps a rudimentary vocabulary and a limited grasp of grammar and syntax¹⁸. Others working in overtly 'people professions', such as social work, will acquire as part of their disciplinary expertise a capacity to assess performances in these respects. Workplace assessors are likely to have a limited appreciation of assessment principles and practice, but set against that do have an appreciation (which may range from the tacit to the codified) of the judgement of performances against organisational expectations.

On some programmes, an overall grade is awarded. More often than not, this involves the averaging of 'percentages' as a first resort¹⁹: less often, other grading scales are used and the overall award determined with reference to the profile of the grades awarded for the programme components²⁰. Whereas it might be possible to grade academic work quite finely, the grading of professional achievements is limited in practice to – at most – a small number of assessment categories. Determining an overall grade from a 'mixed bag' of grades is clearly problematic: it is one of the arguments against the classification of honours degrees in the UK. Profiles of achievement (which can contain both quantitative and qualitative evidence) are likely to be more valid and meaningful.

In a study of assessment of work-based learning in foundation degrees, Woolf and Yorke (2010) found that, whereas academics were content to grade students' work (using various scales), employers were reluctant to go beyond a pass/fail dichotomous assessment. Determining the reasons for limiting grading was not part of the study, but the reasons could include lack of expertise in respect of judging performance against multiple levels; possible role conflict in respect of support and assessment; and other social pressures arising from collegiality (even if only temporary). A different perspective is offered by Willis (forthcoming) who notes that professional training frameworks at the University of Surrey 'require' employer assessors to distinguish different qualities of performance above the satisfactory. There is however no consensus on how many categories for the quality of performance there should be.

The education of assessors

Following Knight (2002) who had argued that summative assessment was in disarray, James (2003:197) observed that:

'Assessment is one of the least sophisticated aspects of university teaching and learning.'

Taken in the round, the quotation is, if anything, a little generous – though there are pockets of sophistication in assessment, notably in the area of medicine. Assessment regulations in higher education institutions across the UK are varied, with little by way of rationale for the choices that different institutions have made (Yorke *et al*, 2008). Assessment practices vary within and between modules, with relatively little attention being given to providing a programme-wide focus on what the students are expected to achieve: the sum is sometimes an incoherent congeries of the parts. This is not the place to bewail weaknesses in summative assessment: the determined reader can consult Yorke (2008) for a detailed argument on the point.

The *Dearing Report* in the UK (NCIHE, 1997) identified the 'assessment problem' but not the scale of it. Innovations in assessment practice tend to be piecemeal, and lacking in strategic or programmatic coherence (in effect, this is another way of stating Knight's, 2002, point about summative assessment being in disarray). There is a general reluctance to lift up the assessment stones to reveal the inappropriate practices that have existed peacefully there for some considerable time (Yorke, 2008). Much that goes on in assessment passes relatively

¹⁸ The assessment of achievements in the subject discipline is more problematic than many acknowledge (see Yorke, 2008 for an extended argument), but that issue is being 'bracketed out' of the present discussion.

¹⁹ To call marks 'percentages' is actually erroneous, since in the vast majority of cases there is no descriptor of what 100% might mean. Exceptions occur in subjects where it is possible to obtain a perfect score, such as when responding to computer-generated questions.

²⁰ Quite often, in the case of averaging, there is a secondary recourse to profiling if the mean percentage falls into a borderline category. The student is awarded the better of the two computed overall grades.

Assessing the complexity of professional achievement : Learning to be Professional through a Higher Education e-book
<http://learningtobeprofessional.pbworks.com/>

unquestioned – after all, if nothing happens to disturb the even tenor of assessment’s way, there is little stimulus to review practices.

This might be a reasonable interpretation of a study by Knight and Page (2007) of the assessment of ‘wicked competences’ in a number of broadly vocational subjects. ‘Wicked competences’ are described as

achievements that cannot be neatly pre-specified, take time to develop and resist measurement-based approaches to assessment. (Knight and Page, 2007:2)

Knight and Page used the following as ‘wicked competences’ (though one might take issue regarding the use of the term ‘competences’ for all of them, it is the implications for assessment that are of interest here²¹):

- Developing supportive relationships
- Emotional intelligence
- Group work
- Listening and assimilating
- Oral communication
- Professional subject knowledge
- Relating to clients
- Self-management (confidence and effectiveness)
- ‘Taking it onwards’ – acting on diagnoses

An on-line survey of 83 ‘key informants’ found that they seemed generally to hold the view that the assessment of the set of ‘wicked’ competences was not particularly problematic. The informants may have been sufficiently confident in their experience and expertise to be able to make adjustments for the inevitable variation across their student cohorts. They may have been towards the upper end of the Dreyfus brothers’ categorisation.

None of the competences cited by Knight and Page is, however, straightforward to assess since each subsumes a range of achievements. Whereas the academic who is proficient or expert (in the Dreyfus brothers’ terms) in assessing professional practice might cope well with making judgements of students’ achievements, attention needs to be given to developing the judgemental capabilities of those with lesser expertise²². Workshops involving video-clips of performance, or role-playing, have potential in this respect.

Is there a need to change the summative assessment paradigm?

Though few in higher education would acknowledge it, summative assessment is, as Knight (2002) argued, in disarray. With academic achievements in mind, Sadler (2005) suggested that work could be graded according to the extent to which it met primary and secondary criteria (with greater weight being given to the former). Sadler himself recognised that not all criteria lent themselves to a dichotomous judgement (met *versus* not met), unless they could be treated more loosely in terms of satisfactory *versus* not satisfactory. Sadler’s main argument, for grading against primary and secondary criteria, runs into trouble when the practicalities are considered. Student achievements do not always fit the approach – for example, where secondary criteria are well achieved but the primary ones are weakly achieved. Where the overall judgement is in pass/fail terms, the problems are of little

²¹ Greene’s (2004:5) ‘eight general capabilities’ of highly effective people might have provided a better starting-point for a study of this kind. One would not necessarily expect inexperienced people to possess these capabilities to a high degree, but they do constitute a potential framework within which professional development can be encouraged.

²² The point applies broadly across higher education, given James’ (2003) stricture noted above: however, the focus of this chapter is narrowly on the assessment of professional behaviour.

account but, where some form of scaling of the achievements is concerned (as is so often the case in academic work), the rules by which the component achievements are cumulated become formidably complex.

As is apparent from the general thrust of this chapter, the assessment of professional achievement (in most circumstances) demands a responsive approach. Broad judgements can be made against criteria: most assessors would probably not find difficulty in deciding whether or not a person's achievements met a threshold level, even if they might find it more challenging to elaborate on the rationale for the judgement. (The threshold level has to be determined with reference to contextual factors: the risk of a 'false positive' – i.e. an unmerited pass – has to be lower where safety to the public is a concern than where it is not.) If grading is to be employed, it is only feasible to grade in terms of broad bands. Even then, experienced judges are likely to disagree to some extent about the grade that should be awarded – there are many reports in the literature of disagreements amongst markers of academic work even when the task is more tightly specified than is typically the case in professional practice.

Assessment in higher education is dominated by a paradigm in which 'measurement' looms large. The development of the Higher Education Achievement Report [HEAR], in which graded achievements are supported by qualitative evidence²³, hints at an alternative paradigm. That alternative is more overtly judgemental in character, and might most appropriately be manifested in an approach in which the satisfactoriness or otherwise of achievement is the primary judgement, and the judgement is backed up by qualitative evidence from actual practice. Such an approach would seem eminently suitable for the professional milieu²⁴. With the boundaries between academic and work-engaged learning becoming steadily more blurred, the approach would seem to have a wider applicability.

Key points

The assessment of professional achievement has to be, in large measure, responsive and judgemental, rather than prescriptive and compliance-oriented. The philosophical underpinning is relativism rather than realism.

Consequences of a relativist paradigm of assessment include:

- Assessment is more holistic than where learning outcomes are tightly specified in curricula.
- Grading is problematic. Whilst a criterion of 'satisfactoriness' might be employed, and might be expanded to cover a limited number of grades, fine discriminations become difficult because of the multiplicity of sub-criteria that apply in many cases, and the problem of weighting them. Qualitative evidence might be used to buttress dichotomous (satisfactory/unsatisfactory) judgements.
- The technical characteristics of assessment, whose origin lies in psychometrics, do not generally suit the assessment of professional achievement.
- There is a need for developmental work on the part of assessors in higher education in order to enhance their expertise as assessors, especially where they are relatively inexperienced in dealing with performances in the workplace. There are similar needs in respect of workplace assessors, but meeting these (especially in small and medium-sized enterprises, where resources are limited) is particularly challenging.

References

AUQA (2009) *Setting and monitoring academic standards for Australian higher education* (Discussion paper). At www.auqa.edu.au/qualityenhancement/academicstandards/discussion-paper.pdf .

²³ And, where appropriate, quantitative evidence.

²⁴ See Yorke (forthcoming) for a fuller discussion of the need for a paradigm shift in assessment.

Assessing the complexity of professional achievement : Learning to be Professional through a Higher Education e-book
<http://learningtobeprofessional.pbworks.com/>

- Baxter Magolda, M. (2009) Educating students for self-authorship: learning partnerships to achieve complex outcomes. In C. Kreber (ed) *The university and its disciplines: teaching and learning within and beyond disciplinary boundaries*. Abingdon: Routledge, 143-156.
- Belbin, R.M. (1981) *Management teams: why they succeed or fail*. London: Heinemann.
- Billett, S. (2009) Realising the educational worth of integrating work experiences in higher education. *Studies in Higher Education* 34 (7), 827-843.
- Cairns, L. and Stephenson, J. (2009) *Capable workplace learning*. Rotterdam: Sense Publishers.
- Dreyfus, H.L. and Dreyfus, S.E. (2005) Expertise in real world contexts. *Organization Studies* 26 (5), 779-792.
- Eisner, E.W. (1985) *The art of educational evaluation*. Lewes: Falmer
- Eraut, M. (2004) A wider perspective on assessment. *Medical Education* 38 (8), 803-804.
- Eraut, M. (2009) How professionals learn through work. Chapter A2 in N, J. Jackson (ed) *Learning to be professional through a higher education e-book*. At <http://learningtobeprofessional.pbworks.com/w/page/15914952/How-professionals-learn-through-work> .
- Eraut, M. (2010) The balance between communities and personal agency: transferring and integrating knowledge and know-how between different communities and contexts. In N Jackson and R Law (eds) *Enabling a more complete education: encouraging, recognizing and valuing life-wide learning in higher education*, 45-65. Accessed via <http://lifewidelearningconference.pbworks.com/f/Lifewide+learning+Conference+E-Proceedings+version+April+09+2010.pdf> .
- Goleman, D. (1996) *Emotional Intelligence*. London: Bloomsbury.
- Greene, T.R. (2004) 32 capabilities of highly effective people in any field. At www.scribd.com/doc/2162334/32-Capabilities-of-Highly-Effective-Persons .
- Guba, E.G. and Lincoln, Y.S. (1989) *Fourth generation evaluation*. London: Sage.
- Hays, R., et al (1994) Methods of assessment in recertification. In D. Newble, B. Jolly and R.E. Wakeford (eds) *The certification and recertification of doctors: issues in the assessment of clinical competence*. Cambridge: Cambridge University Press, 187-200.
- Hedlund, J. and Sternberg, R.J. (2000) Too many intelligences? In R. Bar-On and J.D.A. Parker (eds) *The handbook of emotional intelligence: theory, development, assessment, and application at home, school, and in the workplace*. San Francisco, CA: Jossey-Bass, 136-167.
- Hodges, B. et al (1999) OSCE checklists do not capture increasing levels of expertise. *Academic Medicine* 74 (10), 1129-1134.
- Jackson, N. J. (2010a) From a Curriculum that Integrates Work to a Curriculum that Integrates Life: Changing a University's Conceptions of Curriculum *Higher Education Research and Development Work Integrated Learning special edition* 29:5 available on-line at <http://lifewideeducation.pbworks.com/f/HERDSA+PAPER+MAY+2010.pdf> Accessed 23/02/11
- Jackson, N. J. (2010b) Developing professional capability through lifewide education with a specific focus on students' creative development. In N. J. Jackson (ed) *Learning to be Professional through a Higher Education e-book* Accessed via <http://learningtobeprofessional.pbworks.com/w/page/15914981/Learning>
- James, R. (2003) Academic standards and the assessment of student learning: some current issues in Australian higher education. *Tertiary Education and Management* 9 (3), 187-198.
- Jessup, G. (1991) *Outcomes: NVQs and the emerging model of education and training*. London: Falmer.
- King, P.M. and Kitchener, K.S. (1994) *Developing reflective judgment: understanding and promoting intellectual growth and critical thinking in adolescents and young adults*. San Francisco: Jossey-Bass.

- Knight, P.T. (2002) Summative assessment in higher education: practices in disarray. *Studies in Higher Education* 27 (3), pp.275-286.
- Knight, P. (2005) *Assessment and complex learning*. At <http://kn.open.ac.uk/public/document.cfm?docid=6576> .
- Knight, P. and Page, A. (2007) *The assessment of 'wicked' competences [Report to the Practice-based Professional Learning Centre]*. At www.open.ac.uk/cetl-workspace/cetlcontent/documents/460d21bd645f8.pdf
- Kohlberg, L. (1964) *The philosophy of moral development: moral stages and the idea of justice*. San Francisco: Harper and Row.
- Lester, S. and Costley, C. (2010) Work-based learning at higher education level: value, practice and critique. *Studies in Higher Education* 35 (5), 561-575.
- Mager, R.F. (1962) *Preparing instructional objectives*. Palo Alto, CA: Fearon.
- Mayer, E. (Chair) (1992) *Key competences*. Melbourne: Australian Education Council and Ministers of Vocational Education, Employment and Training.
- NCIHE (1997) *Higher education in the learning society (Report of the National Committee of Inquiry into Higher Education: 'The Dearing Report')*. Norwich: HMSO.
- Perry, W.G. (1998) *Forms of ethical and intellectual development in the college years*. San Francisco: Jossey-Bass [Reprint of 1970 text, with a new introduction by L.L. Knefelkamp].
- Sadler, D.R. (2005) Interpretations of criteria-based assessment and grading in higher education. *Assessment and Evaluation in Higher Education* 30 (2), 176-194.
- Salovey, P. and Mayer, J.D. (1990) Emotional intelligence. *Imagination, Cognition, and Personality* 9 (3), 185-211.
- Schuwirth, L.W.T., et al (2002) When enough is enough: a conceptual basis for fair and defensible practice performance assessment. *Medical Education* 36 (10), 925-930.
- Simon, H.A. (1957) *Models of man*. New York: Wiley.
- Stephenson, J. (1992) Capability and quality in higher education. In J Stephenson and S Weil (eds) *Quality in learning: a capability approach in higher education*. London: Kogan Page, pp.1-9.
- Sternberg, R.J. (1997) *Successful intelligence: how practical and creative intelligence determine success in life*. New York: Plume.
- Stowell, M. (2004) Equity, justice and standards: assessment decision making in higher education. *Assessment and Evaluation in Higher Education* 29 (4), pp.495-510.
- Willis, J., Sahama, T. and Hargreaves, M. (2010) Assessing Performance and Capability in Work Placements: A collaborative study involving Queensland University of Technology Australia and the University of Surrey England. In N. J. Jackson (ed) *Learning to be Professional through a Higher Education e-book*. Accessed via <http://learningtobeprofessional.pbworks.com/w/page/15914981/Learning> .
- Willis, J. (forthcoming) Developing Professional Capability at the University of Surrey. In N. J. Jackson (ed) *Learning to be Professional through a Higher Education e-book*
- Woolf, H. and Yorke, M. (2010) *Guidance for the assessment of work-based learning in foundation degrees*. Lichfield: Foundation Degree Forward. At [www.fdf.ac.uk/downloads/290/20101004145843SACWG Assessment report.pdf](http://www.fdf.ac.uk/downloads/290/20101004145843SACWG%20Assessment%20report.pdf)
- Yorke, M. (1998) Assessing capability. In J Stephenson and M Yorke (eds) *Capability and quality in higher education*. London: Kogan Page, pp.174-191.
- Yorke, M. (2003) Going with the flow? First cycle higher education in a lifelong learning context. *Tertiary Education and Management* 9 (2), pp.117-130.

- Yorke, M. (2008) *Grading student achievement: signals and shortcomings*. Abingdon: Routledge.
- Yorke, M. (2011) Summative assessment: dealing with the 'measurement fallacy'. *Studies in Higher Education* (in press).
- Yorke, M. (forthcoming) Work-engaged learning: towards a paradigm shift in assessment. *Quality in Higher Education*.
- Yorke, M. and Knight, P. (2006) *Embedding employability into the curriculum*. York: The Higher Education Academy. Also at www.heacademy.ac.uk/assets/York/documents/ourwork/employability/id460_embedding_employability_into_the_curriculum_338.pdf
- Yorke, M. *et al* (2008) Enigmatic variations: honours degree assessment regulations in the UK. *Higher Education Quarterly* 62 (3), pp.157-180.