A review of vocational research in the UK 2002-2006: Measurement and accessibility issues

Martin Johnson

Abstract

The UK vocational assessment landscape is changing and research has a key role in helping to identify trends and anticipate emerging issues and challenges. This review collates the messages from recently published research about specific issues that have been found to affect the measurement of vocational learning. It also highlights issues that have been found to influence learner access to vocational qualifications. The data collection method was a structured literature review of material published since 2002. The review used three different types of search: a search of articles found in a selected group of internationally respected peer-reviewed academic journals; an online database search; and a search of articles written by respected authors working in the vocational assessment field.

Introduction: Scope of the review

The assessment landscape in the UK is changing. The government is committed to creating more flexible learner pathways involving, amongst other things, the establishment of a structure that equates the relative values of different qualifications with each other. According to the UK examinations and qualifications regulator, the Qualifications and Curriculum Authority (QCA), the National Qualifications Framework (NQF) for England, Wales and Northern Ireland sets out the levels at which qualifications are recognised, helping learners to make informed decisions about the qualifications they need. The framework enables comparisons to be made about the levels of different qualifications and the identification of clear progression routes. Currently there are moves to

reform the NQF in order to develop a structure that allows for the accumulation and transfer of credit achievement over time. This new framework is intended to be unit-based and supported by a system of credit accumulation and transfer, with the achievement of learners expressed through credit, indicating the size of the unit or qualification, and by level, indicating the challenge.

This review was undertaken to gain a sense of the contemporary research messages relating to measurement and accessibility factors currently affecting vocational qualifications. This rationale is situated within the shifting policy context outlined above, where changes are often conducted after consultation processes involving key stakeholders. Underpinning the review is the notion that such consultations should be informed by messages from educational research which can help to facilitate reasoned debate. It also aims to gain a sense of the extent of educational research in vocational qualifications and whether any areas lack coverage, potentially promoting the strategic planning of future research decisions.

Whilst surveying material for the review it became clear that Australia is a rich source of recent research literature, dealing particularly with aspects of measurement such as reliability and validity (e.g. Williams & Bateman, 2003; Queensland Government, 2005; Clayton et al., 2004; Thomson et al., 1996; Gillis & Bateman, 1999). This reflects the significant developments witnessed by the Australian vocational education and training (VET) sector over the last few years. Parallel to the creation of the NQF in parts of the UK, the development of the Australian Qualifications Framework (AQF) gives added weight to the suggestion that VET principles and practice sometimes converge across national contexts (Misko, 2006). It also helps to explain the presence in some Australian literature which references contemporary developments in the UK (e.g. Clayton et al.; 2004; Thomson et al., 1996; Keating et al.; 2002; Turner, 2002). In a similar spirit, this review hopes to contribute to wider knowledge by drawing the reader's attention to some of the issues concerning research in vocational assessment in the UK.

The review is structured around the two key themes of measurement and accessibility, with the measurement section subdivided separately into reliability and validity issues. After outlining the method used for the review, the findings are arranged into three broad parts: *Measurement: validity; Measurement: reliability;* and *Accessibility.* Because these sections are lengthy, a final discussion of the key points raised is included before ultimate conclusions are made.

Method

An important element of scientific review methodology is its potential replication. It is worth noting that scientific systematic review methodologies developed in the late 1990s in response to critiques of social science which suggested that the qualitative descriptive methods used in social science lacked evaluative use. Systematic scientific review methods were particularly favoured by medical practitioners seeking to promote evidence based practices in health care provision through evaluating the efficacy of different intervention studies. Evans and Benefield (2001) state that the key features of a systematic review are: the formulation of an explicit research question; transparency of method; exhaustive searches that seek both published and unpublished studies; clear criteria to assess that quality of studies; clear criteria for excluding studies based on the scope of the review; joint reviewing to reduce potential bias; and a clear statement of the findings. Based on these criteria this study could be said to have adopted a semi-systematic approach. We chose not to employ an explicit strategy to evaluate the weight of the findings for each study in the review since this aspect of the systematic review method reflects the roots of the method in evaluations of efficacy in medical intervention studies. Some commentators have reservations about whether the systematic review method can translate from the medical context to other domains. Evans and Benefield (2001) state that the method can reduce questions of research to those of technical efficiency and effectiveness and discourage research exploring 'the wider social, philosophical or ethical issues which are implicit in social policy decisions' (Evans and Benefield, 2001, p. 539). Boaz et al. (2002) also argue that systematic reviews tend to prioritise certain methods over others, particularly randomised controlled trials, and can overlook the impact of context that often leads to multiple outcomes.

We began the review by selecting a suitable definition of VET. VET has been conceptualised in a variety of ways, often differentiating generic employability skills that raise awareness of the world of work from specific job-related competencies required for a career in a particular vocational sector. It is common to encounter dichotomous labels which distinguish between 'vocationally related'/'pre-vocational' and 'vocational learning'. The terms 'weakly vocational' and 'strongly vocational' (Stanton, 2004) have been used to perform a similar function.

Stasz & Wright adopt a broad definition of vocational learning. They define it as:

> Those experiences and activities that lead to (or aim to lead to) outcomes which are formally designated, through everyday discourse at both policy and institutional levels, as being vocational, whether such outcomes are

in the form of vocational qualifications…or, where the learning is not certificated, the intended outcomes are clearly linked with developing an individual's competence in his/her current occupational area. (Stasz & Wright, 2004, p.3)

We based our first inclusion criteria on this definition since it accommodates issues in assessment that are both certification-related and more general. The selection process was refined using two additional criteria. Literature was included that dealt with either: measurement issues (e.g. reference to reliability or validity); accessibility issues (e.g. reference to factors that potentially affect access to learning); or qualifications offered for 14-19 year olds in the UK (e.g. NVQ, VCE, AVCE, vocational GCSE, Key Skills, Skills Test, and Vocationally-related Certificate).

It was necessary to establish the publication date margins for incorporation and the keywords that would structure the investigation. The four year review period was chosen for both practical and principled reasons. Balancing the project resources with the desire to investigate 'contemporary' research messages led to a consideration of evidence dating from 2002. It is worth noting that some of the literature contained in the reference list predates 2002. This is because some older articles offer contextual evidence (for descriptive purposes) or they contain issues which are referred to in the contemporary publications. A set of keywords were chosen by the researchers to cover the scope of the investigation. The words chosen were: vocation/al; qualifications; valid/ity; reliable (ility); fair/ness; access/ibility; measure/ment; NVQ; VCE; AVCE; vocational GCSE; Key Skills; Skills Test; and Vocationally-related Certificate.

The review entailed three different search types: a journal content search; an online database search; and a search of known authors respected for their work in the vocational assessment field. The choice of journals, databases and authorities were informed by researchers' experience of the most valuable sources of potential data. These were either assessment, sectoral or domain specific searches. In total, 86 relevant articles were identified.

Journal search

The contents pages of a selected group of internationally respected peer-reviewed academic journals were searched. These were: *Assessment in Education: Principles Policy and Practice; the British Educational Research Journal; Education and Training; Educational Research; the Journal of Education and Work; the Journal of Further and Higher Education; the Journal of Vocational Education and Training; and Research in Post-Compulsory Education.* This search yielded 43 articles.

Database search

The results of the journal search were supplemented by a search of eight online databases. These were: the *Nuffield 14-19 Review of Education and Training* database; the database of the *British Education Index; CEDEFOP* collection (the European Centre for the Development of Vocational Education and Training); *Learning and Skills Research Centre* publications archive; *ISI Web of Knowledge* portal; *International Bibliography of the Social Sciences* database; the *British Library's Electronic Table of Contents* catalogue; and *IngentaConnect.* This allowed access to a wider variety of documents, including articles in less well known peer-reviewed journals. It also supplied references to an assortment of conference, working and discussion papers. 32 articles were found through this search.

Author search

Seven key UK-based academic authorities on vocationally related assessment were identified. These were: four university professors working in different Schools of Education; two university lecturer/researchers from different Schools of Education; and the head of research in a large governmental education agency. In this search we identified eleven documents not covered elsewhere in the review. The balance of these texts, largely comprising either peer reviewed journal articles, conference papers or government department funded research papers, is weighted towards accessibility issues.

Analysis

The literature identified in the three searches were read in detail and analysed. Coverage of accessibility issues outweighed that of measurement issues.

Criterion	Journal articles	Conference papers	Government and exam body reports	News articles	
Measurement	9	4	9		22
Accessibility	13	6	8	4	31
14-19 qualification	24	2	6	1	33
	46	12	23	5	86

Within the broad themes of measurement and accessibility a number of salient themes emerged:

Measurement - Validity	purpose of learning and assessment
	language use and performance descriptors
	fairness
	work-based learning and assessment
Measurement - Reliability	assessor skill and experience
	centralisation and flexibility
	different assessment methods
	equal opportunities and variability
Accessibility	learning culture and environmental obstacles
	vocational identities
	vocational value
	academic demand
	parity of esteem

Measurement - Validity

To begin with, a broad definition of validity was identified in the literature. Messick states that 'validity is an integrated evaluative judgement of the degree to which empirical evidence and theoretical rationales support the adequacy and appropriateness of inferences and actions based on test scores or other modes of assessment' (Messick, 1989, p. 13). This broad definition was adopted for the purposes of this review since it subsumes other validity categories, such as construct, external, inferential, consequential and content validity.

Purpose of learning and assessment

In a conceptual piece written for the UK Learning and Skills Research Centre, Stasz & Wright (2004) suggest that the target culture underpinning UK government policy development assumes that qualifications can play an important systemic function. Data about qualification uptake and performance is used to evaluate the functioning of vocational education and training. Stasz & Wright (2004) argue that the assessment process is perceived by policy makers as carrying a great deal of reliable information about the quality of the system as a whole. Consequently, the validity of qualifications is threatened because the sector is dominated by a focus on end-points rather than concerns about the learning process. Moreover, in a literature review, Stasz et al. (2004) go on to argue that a conflation of different agendas within government also potentially

undermines the validity of the vocational offer. Attempts to introduce young people to occupational options, up skill the national workforce, and court the social inclusion agenda can lead to a 'vague vocationalism'. We might infer that this represents a shift of emphasis. Instead of valuing vocationally related learning for its direct contribution to particular skills development in particular vocational contexts there is a greater focus on the motivational qualities of vaguely work-related learning. The measure of success for such courses lies primarily in their ability to counter disaffection rather than their positive contribution to specific skill sets or the continued inspiration of already motivated individuals.

Chelbi (2006) and Tolley et al. (2003) highlight another set of conflicting purposes in vocational education that potentially undermine the validity of the inferences made from assessment data. They suggest that diverging roles (e.g. trainer, learner or employer) or sector related perspectives can lead to a multiplicity of expectations about what constitutes valid vocational learning. This might also explain why some stakeholder groups place less value on the final qualification (McAdam & Crowe, 2004). Tolley et al. (2003) contrast employers' expectations (e.g. training should apply to the particular context of their workplace) with those of college trainers' seeking to encourage learner sector readiness and a broader experience.

Language use and performance descriptors

In a theoretical critique of 'outcomes approaches' in vocational qualifications, Oates (2004) cautions that attempting to generalise performances from context-bound situations can potentially compromise the inferential validity of an assessment. Efforts to define 'competence' in a generic sense have tended to result in a series of partial technical descriptions, making inference beyond the immediate context difficult. In a conceptual piece, Hager (2004) explains that generalising requires competency statements to employ linguistic abstractions, deliberately omitting a myriad of possible features that could have been included in any particular situation. A number of empirical studies report concerns that written performance descriptors might be a source of invalid assessment judgements. These include: inaccessible language forms (Swailes & Roodhouse, 2003); learner dependence on others to interpret the assessment standards and evidence collection (Tolley et al., 2003); and misaligned dispositions held by different actors within the vocational learning process, leading to a discrepant regard for different values (e.g. candidates' personal attributes rather than their knowledge or skills) (James & Diment, 2003; Yaphe & Street 2003). According to Bourdieu, dispositions are part of the schemata that prefigure individuals' actions. They are influenced by social interactions and environmental factors which in turn affect practices and perceptions (Bourdieu and Wacquant, 1992).

There are suggestions that competence is more likely to be inferred from observing how an individual performs on a number of occasions and in a number of settings. Using case study evidence, Fuller et al. (2003) suggest that this might involve the interaction of both tacit (practical) and codified (explicit) knowledge. Consequently, experts might need to have a key role in the assessment process, inferring those tacit rules that underlie observable behaviours.

This also has implications for the external validity of qualifications. Oates (2004) cites evidence that employers increasingly use informal information about candidates' prior employment experiences in their selection decisions. Reporting findings from interviews with senior human resources staff, Wolf & Jenkins (2002) also highlight the developing trend for employers to value alternative sources of information beyond qualifications when recruiting new employees (e.g. data from psychometric tests).

Fairness

Critiquing outcomes-focused methods (e.g. Jessup, 1991), Oates (2004) argues that the language used for defining and assessing competence tends to be partial and prone to reinforcing inequity in work-based learning programmes. He maintains that the language used to define standards and assess performance is value-laden and partly responsible for the cultural transmission of dominant attitudes and possibly replicating inequalities. Performance descriptions in occupational standards can exclude certain forms of skill (e.g. certain communication styles where preferences tend to be particularly gender-related). Taking a critical perspective of the language that defines competencies can help to remove structural obstacles to equal opportunities.

Similarly, discussing the challenges surrounding equal opportunities in vocational settings, Murphy et al. (2003) argue that vocational learning often replicates the structural barriers that exist in the wider world of work. They argue that systematic diversity makes fairness difficult to conceptualise. They list a number of features (the existence of multiple qualification awarding bodies, the diversity of training and workplace settings, and the multiplicity of assessors and verifiers) which combine to make the monitoring of learning provision quality and equity difficult to achieve. Some models focus on eliminating deliberate, unfair discrimination at the point where entitlements and services become available to individuals. Murphy et al. (2003) argue that inequality needs to be understood as a structural problem. Social factors militate against some groups reaching the point of learning access where safeguards against unfair treatment might exist. This helps to explain how unequal employment opportunities and workplace practices might continue despite the

existence of equal opportunities strategies. Murphy et al. (2003) also discuss how grading might perpetuate inequalities. They suggest that the generation of more detailed performance descriptors, although helping to discriminate between outcomes, risks incorporating factors that are not completely valid. They argue that construct irrelevant aspects of performance, such as perhaps the quality of presentation, might be in accordance with an assessor's cultural dispositions and might advantage particular learner subgroups.

Work-based learning and assessment

Situated cognition theories imply that intellectual competencies are heavily context dependant (Griffiths & Guile, 2003), and that work-based assessments might lack generalisibility. Boreham et al. (2002) suggest that the concept of 'work process knowledge' can help to explain this through highlighting the interdependent relationship between learning and context. Work process knowledge involves workers having an expanded understanding of their role in the context of the organisation within which they operate. Crucially, this knowledge is constructed in the process of work when solving problems that arise from inevitable tensions between the worker's activities as learned and the reality experienced during work. Furthermore, work process knowledge is assumed to be held collectively rather than individually, resulting in a workplace culture.

Tolley et al. (2003) acknowledge that the authenticity of work place problem solving situations contributes positively to learner capability and motivation. They contrast this with some college-based simulated work settings which lacked the authentic practices and contingencies of real environments. They suggest that this might encourage limited transfer between college and workplace.

Measurement - Reliability

Reliability has been defined in a number of ways. For Gillis & Bateman (1999) it refers to the extent to which the assessment can theoretically provide repeatable outcomes for candidates of equal competence at different times and/or places. Others suggest that reliability needs to be defined as an aspect of validity. Feldt & Brennan state that 'reliability refers to a concept much closer to the measurement concept of validity' (1989, p.106). In this sense, reliability concerns the validity of inferences about the stability or repeatability of scores/grades across time, raters, or tasks. This is an important consideration since low levels of judgement consistency or reliability undermine the potential for mutual recognition between different qualifications. This is because mutual recognition requires widespread confidence in the consistency of assessment decisions

across contexts. Most of the reliability literature that we found focuses on the challenges posed by the interaction of differing contexts and social mediation in the assessment process.

Assessor skill and experience

Reviewing different qualifications models, Clayton et al. (2004) identify a number of threats to reliability at systemic and individual levels. At an individual level, assessment against competency standards requires evidence collection and the formation of a judgement about achievement. Clayton et al. (2004) suggest that assessors' decision making processes might be influenced by their level of skill and experience. The involvement of assessors' tacit knowledge of a particular technical domain is an important aspect of vocational assessment. This might be problematic. Models of competency might not be consistent across different assessors, leading to differential judgements about performance quality. This effect can be mitigated when assessors work in a community of practice (Greatorex & Shannon, 2003). Clayton et al.(2004) note that subject specialism does not necessarily equate with assessment specialism.

In order to overcome this potential threat to consistency, Clayton et al. (2004) suggest that systemic checks and balances need to be instituted. They argue that increased emphasis on 'front end' quality assurance procedures could alleviate some of the reliability problems through recognition and accreditation processes. Ongoing professional development for assessors would help them to share tools and practices, communicate advice about gathering and evaluating evidence, as well as instituting a common dialogue across contexts.

They also argue that quality control 'feedback' procedures can increase systemic centralisation. Monitoring the acceptability of assessment outcomes after the event might have a knock on effect on assessment validity. They cite literature which suggests that standardised moderation exercises, identifying inaccurate and inconsistent applications of standards, will tend to concentrate on easily measurable statements that can be moderated across different learning contexts.

Centralisation and flexibility

This tension between centralised and devolved elements also affects assessment tools. Torrance & Coultas' (2003) review of over 100 books, articles and research reports suggests that learners prefer coursework and practical competence-oriented assessment over end-of-course tests. Discussing current research about the assessment of generic skills, Curtis (2004) suggests that work-based assessments present problems for reliability measures, despite centralised training and moderation procedures. Whilst different raters can be trained to

consistently judge a complex task, the real hurdle is task-sampling variability. The potentially large range of tasks makes it difficult for assessors to be trained to assess them all reliably. Task-variability means that a number of tasks need to be assessed in order to achieve satisfactory levels of reliability. High costs also mean that learners might only complete relatively few tasks of this nature. It is also possible that some learners might be assessed on more or less favourable tasks than others.

Different assessment methods

Curtis (2004) reviews the strengths and weaknesses of different assessment methods. He suggests that holistic judgements, although less reliable in social or workplace settings, can work reliably in schools because of frequent and close observations of learner performance. Findings from a research study by Greatorex & Shannon (2003) suggest that the use of holistic judgements makes it inappropriate to apply traditional concepts of reliability to such assessments. This is because holistic judgements cannot be treated as individual items for statistical analysis.

Portfolios can provide useful evidence of developing processes and can claim a good degree of validity, but they are not readily digestible or comparable. Workplace assessment, whilst being highly authentic and valid, can resist standardisation, making comparisons across contexts difficult.

Equal opportunities and variability

Murphy et al. (2003) argue that the availability of arrangements for candidates with special requirements presents important barriers to the reproducibility of assessments. Equal opportunities considerations require a degree of human judgement in order to balance the needs of the candidate against those of the assessment procedures. They go on to state that the 'Equal opportunities discourse clearly severs the notion of reliability from mathematical models of reproducibility' (2003, p. 12), suggesting that the application of human judgement in a social context affects the consistency of final assessment judgements.

Accessibility

The issue of access has affected developments in the UK vocational sector since the 1980s. The Review of Vocational Qualifications in England and Wales Working Group (De Ville, 1986) acknowledged a number of structural weaknesses. This review also identifies a number of key issues which might

affect learners' access to vocational learning. Some of these relate to tensions between academic and vocational learning paths, and the degree to which these interact with the aspirations of learners. Others concern structural obstacles to learning access at individual and system level.

Learning culture and environmental obstacles

The learning landscape of 14-16 year olds in the UK is currently shifting. Smith (2005); Marsh (2005) and Whittaker (2005) report the expansion in the number of part-time vocational programmes now available to school aged learners. One news report (Whittaker, 2005) also suggests that the number of 14-16 year olds in further education and training will double in the next eight years.

Findings from the Teaching and Learning Research Programme, (e.g. Hodkinson et al., 2004; Davies & Biesta, 2004) suggest that 'learning cultures' are an important analytic tool for understanding the experiences of learners. Learning cultures involve the practices through which people learn, allowing a broad analysis of issues that might influence access to vocational learning. A number of studies have suggested that aspects of the learning cultures of schools and colleges, and more particularly their differences, can influence motivation to access vocationally related learning. For example, Davies (2003) reported that school aged students enjoyed the practical involvement of applied skills in a college setting.

Dissonance between school- and college-based learning is harnessed by government initiatives such as the Increased Flexibility Programme (IFP) which brings school aged pupils into post-16 learning environments. Evaluating the IFP programme, Golden et al. (2005) suggest that part of the success of the initiative is its ability to engage learners who were formerly school-refusers. They also found that less able students benefited the most from flexible learning provision. Some research studies suggest that these successes might be linked with pedagogical factors that are associated with the cultural conditions of college learning. Davies (2003) suggests that school aged learners' engagement with college-based learning might be influenced by their tutor's pedagogy. Having a greater degree of learning independence and having a more flexible tutor/learner relationship appeared to engage the learners. Ecclestone (2004) also found that learners working towards vocationally related qualifications valued having relatively higher levels of autonomy and independence in college compared with school.

The inclusion of younger learners in college-based education has potential problems. Media reports (Whittaker, 2006[a]; 2006[b]; 2006[c]) suggest that funding differentials between school and FE college staff might lead to conflict. There are also some concerns about the potential vulnerability of younger learners in such

environments. Research studies looking at cohorts of learners (Davies & Biesta, 2004) and student teachers from post-compulsory education (Partington, 2003) suggest that further education faces a challenge in order to cater for potentially vulnerable non-traditional learners. It is suggested that changing teaching staff dispositions might be a very large and important challenge in order to effect the possible cultural transformation required to support such learners. Reporting on research for the Learning and Skills Development Agency, Duckett (2005) also suggests that teacher dispositions present a potential hurdle to flexible learning programmes. He argues that some teachers still employ deterministic paradigms and that they need to acknowledge that employability skills can be taught and are not linked to learner maturity levels. 'Misalignment' is also alluded to in the research study findings of Davies & Tedder (2003) and Bloomer et al. (2002). They report that the younger learners involved in their studies had limited longer term vocational goals. This contrasts with the dominant assumptions behind current vocational initiatives which suggest that learners are often motivated by enduring learning and career aspirations.

There are a number of other suggestions why vocationally-related learning may have enjoyed limited success in school settings. Maynard & Smith (2004) report on an otherwise successful partnership between a school and a training provider which had to overcome difficult timetabling considerations. Wolf (2002) and Gadd (2004) also highlight that schools generally struggle to keep up to date with the necessary resources to support vocational learning, potentially undermining learning and achievement. This might help to explain why Golden et al. (2005) found that shared training between schools and external providers supported better learning outcomes.

Vocational identities

There is broad agreement in the literature that a number of external pressures can influence vocational learners' identities and their willingness to access vocational learning (Davies & Tedder, 2003; Bloomer et al., 2002; Forsythe & Furlong, 2003; Ecclestone, 2004; Gibson, 2004; Colley, 2003; Raffo, 2003). Case study data (Davies, 2003) illustrates how extrinsic pressures can affect the ability of further education college learners to plan ahead. She argues that students' learning careers tentatively unfold and might be culturally determined. She suggests that the lack of economic and social capital might influence learners' perspectives, making them more comfortable in a work-related learning environment. This is also supported by Bloomer et al. (2002) and Worth (2002) who argue that learner progression routes are the product of constrained choices. Material necessity and limited familial experience of post-16 education can influence their aspirations. The heavy influence of school, family and friends on individual learning choices is also noted in media (Mansell, 2006), evaluation (Golden et al., 2005) and research reports (Forrest, 2005; Conlon, 2002).

Vocational value

Stasz et al. (2004) and Gallacher et al. (2004) cite literature suggesting that successful work-based learning is more likely to be found in learning institutions that value training. Fuller & Unwin (2003[a]; 2003[b]) suggest that companies in sectors where there is a tradition of apprenticeships are more likely to support work-based learning. Empirical evidence from learners suggests that employer support, e.g. allocated time to study, is an important factor in learning progression (Chelbi, 2006; Bowers-Brown & Berry, 2005).

Variable opportunities to access and/or complete learning might be the result of a number of other sector related factors. Research studies looking at different learning contexts) observe that high staff turnover rates and emphasis on short term sales targets in some sectors might leave employers unable to provide sustainable work-based learning programmes (Maynard & Smith, 2004; Pratten & Curtis, 2002; Spielhofer & Sims, 2004b; Swailes & Roodhouse, 2003; Cockrill, 2002). Spielhofer & Sims (2004[b]) also suggest that some employers may not be aware of the relevance of some vocational work-based learning programmes, a point supported by the empirical findings of Allies et al., 2006.

Case study evidence reported by McAdam & Crowe (2004) suggests that institutionalised undermining of vocational qualifications might arise where employers and learners fail to value the qualification procedures. They highlight workplace practices that abuse assessment procedures, attempting to speed up employee qualification completion rates through the widespread copying of portfolios.

Academic demand

A criticism by some employers, found in literature review findings, is that too much formal teaching and assessment can cause learners to drop work-based vocational study (Stasz et al., 2004). Maynard & Smith (2004) also cite literature that younger, lower ability learners can find vocational qualifications difficult to achieve in the workplace.

These arguments suggest a tension between learner or employer expectations and the requirements of formal qualification accreditation. Spielhofer & Sims (2004[a]) report that some employers are uncomfortable with the large degree of external control and regulation implied by accredited vocational qualifications. Stasz et al. cite literature suggesting that some vocational assessment procedures are cumbersome and time consuming. McAdam & Crowe (2004) reinforce this when they argue that assessment procedures can be bureaucratic for candidates and assessors. It is suggested that a possible consequence of this is learner and employer alienation and demotivation. McAdam & Crowe (2004) also suggest

that standards are sometimes couched in complex language, leading verifiers to give sometimes inconsistent advice and undermining the qualification's credibility.

Parity of esteem

Commenting on the landscape of vocational qualifications development, Brockington (2002) observes that system level factors can militate against schools offering vocational courses. He describes the concerns of head teachers and local government officers who suggest that a wider means of reporting and recognising learners' achievement is needed in order to overcome the vocational/academic divide. It is generally perceived that schools tend to value most highly those qualifications that bolster their relative position with other schools as measured by assessment outcomes. Remotivating disengaged learners relies on schools and colleges providing good quality vocational courses.

Discussion

The content profile of the articles included in the review tends to confirm the observations of Unwin et al. (2003) who state that the research-based literature on vocational qualifications is thin, and that the UK lacks a substantive evidence base on vocational qualifications. This review also suggests that there appears to be a relative paucity of work relating to purely measurement issues and vocational qualifications.

To summarise, the main issues and arguments identified in this literature review are as follows:

Measurement - validity

The simultaneous policy goals of introducing young people to occupational options, raising workforce skills at a national level, and engaging disaffected learners might undermine the validity of vocational qualifications.

Performance descriptors are a potential source of invalid inferences. They might value explicit codified behaviours and overlook tacit underlying skills, or incorporate factors that are not completely relevant or valid to notions of competency. This has consequences for the under-researched issue of fairness, which is associated with the relationship between assessment rules and access.

Measurement - reliability

The importance of consistent assessment judgements is considered to be an important issue for the vocational learning sector as a whole. Low levels of judgement consistency or reliability pose a threat to the possibility of mutual recognition between different qualifications. Levels of assessor skill and experience can contribute to variations in assessment outcomes.

There are tensions inherent in the use of standardised assessment tools in a multiplicity of contexts. A variety of different assessment methods exist but they all have particular strengths and weaknesses depending on the context in which they are used.

Human involvement and variability in the assessment process, particularly with regard to arrangements for candidates with special requirements, presents important barriers to the reproducibility of assessments.

Despite these concerns, there are significant gaps in recent reliability research in the UKand it appears that we need to go back some time to find empirical work addressing such issues (e.g. Betteridge et al., 1996; Wolf and Silver, 1986; Wolf and Silver, 1995; Murphy et al., 1995).

Accessibility

Aspects of the learning cultures of schools and colleges, and more particularly their differences, can influence motivation to access vocationally related learning.

A variety of extrinsic pressures can influence learners' willingness to access vocational learning. Learner progression routes might not simply be the product of learner choice but rather of constrained choices, with learners tending to choose courses based on non-academic factors, such as financial, geographical and social considerations.

There are structural barriers to vocational access. The importance of outcomes-based accountability measures can restrict schools from offering vocational courses that can re-motivate pupils and tackle the social inclusion agenda. A de-coupled structure that recognises and reports learners' vocational achievements without equating them with academic qualification scores might undermine the vocational/academic divide.

Conclusion

It is clear that there are coverage issues relating to recent research concerning vocational qualifications in the UK. The imbalance between measurement and access issues in recent UK literature is interesting when considered in the wider context of recent developments to systematically unify the UK qualifications system through linking general and vocational qualifications through the NQF (QCA, 2004). It might be suggested that the apparent lack of measurement related literature is ironic when seen in this context, since there appears to be greater scope for research to explore the technical issues surrounding the construction of a more unified qualifications system. According to the DfES Research Strategy (2006) educational research has a key role in helping agents to understand and anticipate emerging issues and challenges, and the future environment in which they will have to operate. Consequently, it is in the interest of those agencies who are involved with the development of qualification reforms to be aware of the issues from research findings that concern vocational learning and qualifications.

In conclusion it might be argued that policy decisions that affect qualification developments need to be informed by the multiple perspectives captured by research. This can help policy consultation procedures to acknowledge the existence of dissonant positions and help agents to anticipate why policy development is not always as smooth as hoped and outcomes realised as intended.

Martin Johnson is a Research Officer with Cambridge Assessment, the brad name of the University of Cambridge Local Examinations Syndicate and not-for profit organisation.

References

Allies, L, Smith, D and Mazey, R 2006, Developing a business/learning interface: final report (Project Ref: No. RS4306), The Research Centre, City College Norwich.

Betteridge, D, Cundell, S, Kiwan, D and Macleod. V 1996, An Investigation into Internal Verification from Policy to Practice, Centre for Training and Development, University of Brighton.

Bloomer, M, Davies, J & Tedder, M 2002, The pink collar workers: young women, vocational identification and learning, paper presented at the South-West Regional Learning and Skills Research Network Conference, Ilminster.

Boaz, A, Ashby, D and Young, K 2002, Systematic reviews: what have they got to offer evidence based policy and practice? ESRC UK Centre for Evidence Based Policy and Practice, Queen Mary University of London.

Boreham, N, Samurçay, R & Fischer, M, Eds 2002, Work process knowledge, Routledge, London.

Bourdieu, P & Wacquant, L 1992, An invitation to reflexive sociology, Polity Press, Cambridge.

Bowers-Brown, T & Berry, D 2005, Building pathways: apprenticeships as a route to higher education, Education + Training, vol. 47, no. 4, pp. 270-82.

Brockington, D 2002, Key skills within the national qualifications framework: 14-19, (SKOPE Research Paper No. 31), SKOPE Publications, University of Warwick.

Chelbi, N 2006, Employer and student perceptions of Higher Education and Foundation Degrees, The Research and Development Bulletin, vol. 4, no. 1, pp. 50-58, City College Norwich.

Clayton, B, Roy, S, Booth, R & House, R 2004, Maximising confidence in assessment decision-making: Current approaches and future strategies for quality assurance, National Centre for Vocational Education Research, Station Arcade, Australia.

Cockrill, A 2002, The policy implications of a regional case study: skills shortages and provision in the Welsh automotive and electronics industries, Journal of Vocational Education and Training, vol. 54, no. 1, pp.67-84.

Colley, H 2003, Learning to do emotional labour: class, gender and the reform of habitus in the training of nursery nurses, paper presented to the Experiential, Community and Workbased: Researching Learning outside the Academy conference, Glasgow Caledonian University.

Conlon, G 2002, The determinants of undertaking academic and vocational qualifications in the United Kingdom, Centre for Economics of Education, London School of Economics and Political Science, London.

Curtis, D D 2004, 'The assessment of generic skills', in Generic skills in vocational education and training: research findings, Ed J Gibb, National Centre for Vocational Education Research, Station Arcade, Australia.

Davies, J 2003, A taste of further education: the meaning of 'coming to college' for a group of 14-16 year olds, paper presented to the research student symposium of the annual conference of the British Educational Research Association, Edinburgh.

Davies, J & Biesta, G 2004, Coming to college: the experience of vocational learning of two cohorts of 14-16 year-olds in a further education college, paper presented to the annual conference of the British Educational Research Association, Manchester.

Davies, J & Tedder, M 2003, Becoming vocational: insights from two different vocational courses in a further education college, *Journal of Vocational Education and Training*, vol. 55, no. 4, pp. 517-540.

Department for Education and Skills 2006, *DfES Research Strategy*, downloaded from http://www.dfes.gov.uk/research/prospectus/index.cfm 17/1/06.

De Ville, O 1986, *The Review of Vocational Qualifications in England and Wales*, MSC/DES, HMSO London.

Duckett, I 2005, Developing skills for employability, VLSP:16-19 news, , no. 7, summer, Learning and Skills Development Agency, London.

Ecclestone, K 2004, Learning in a comfort zone: cultural and social capital inside an outcome-based assessment regime, Assessment in Education: principles, policy and practice, vol. 11, no. 1, pp. 29-48.

Evans, J & Benefield, P 2001, Systematic reviews of educational research: does the medical model fit? British Educational Research Journal, vol. 27, no. 5, pp. 527-541.

Feldt, L S & Brennan, R L 1989, 'Reliability', in Educational Measurement (3rd ed.), Ed R L Linn, Macmillan, New York.

Forrest, G 2005, Work-related learning at key stage 4: where are we now? VLSP 14-19 news, no. 1, autumn, Learning and Skills Development Agency, London.

Forsyth, A & Furlong, A 2003, Access to higher education and disadvantaged young people, British Educational Research Journal, vol. 29, no. 2, pp. 205-26.

Fuller, A & Unwin, L 2003a, Creating a 'modern apprenticeship': a critique of the UK's multi-sector, social inclusion approach, Journal of Education and Work, vol. 16, no. 1, pp. 5-25.

Fuller, A & Unwin, L 2003b, Learning as apprentices in the contemporary UK workplace: creating and managing expansive and restrictive participation, Journal of Education and Work, vol. 16, no. 4, pp. 407-26.

Fuller, A, Ashton, D, Felstead, A, Unwin, L, Walters, S & Quinn, M 2003, The impact of informal learning at work on business productivity, Final report to the DTI, The Centre for Labour Market Studies, University of Leicester.

Gadd, K 2004, Teaching applied and vocational science, School Science Review, vol. 85, no. 313, pp. 71-8.

Gallacher, J, Whittaker, S, Crossan, B & Mills, V 2004, Modern apprenticeships: improving completion rates, The Scottish Executive, Edinburgh.

Gibson, P 2004, Life and learning in further education: constructing the circumstantial curriculum, Journal of Further and Higher Education, vol. 28, no. 3, pp. 333-46.

Gillis, S & Bateman, A 1999, Assessing in VET: Issues of reliability and validity, review of research, National Centre for Vocational Education Research, Adelaide.

Golden, S, O'Donnell, L, Benton, T & Rudd, P 2005, Evaluation of increased flexibility for 14 to 16 year olds programme: outcomes for the first cohort, National Foundation for Educational Research, Slough.

Greatorex, J & Shannon, M 2003, How can NVQ assessors' judgements be standardised? Paper presented to the annual conference of the British Educational Research Association, Edinburgh.

Griffiths, T & Guile, D 2003, A connective model of learning: the implications for work process knowledge, European Educational Research Journal, vol. 2, no.1, pp. 56-73.

Hager, P 2004, The competence affair, or why vocational education and training urgently needs a new understanding of learning, Journal of Vocational Education and Training, vol. 56, no. 3, pp. 409-34.

Hodkinson, P, Biesta, G & James, D 2004, Towards a cultural theory of college-based learning, paper presented to the annual conference of the British Educational Research Association, Manchester.

Jessup, G 1991, Outcomes: NVQs and the emerging model of education and training, The Falmer Press, London.

James, D & Diment, K 2003, Going underground? Learning and assessment in an ambiguous space, Journal of Vocational Education and Training, vol. 55, no. 4, pp. 407-22.

Keating, J, Medrich, E, Volkoff, V & Perry, J 2002, Comparative study of VET systems: National systems across three regions under pressure of change - Review of research, National Centre for Vocational Education Research, Station Arcade, Australia.

Mansell, W 2006, Still cultivating a split society, The Times Educational Supplement, Friday 24th March.

Marsh, J 2005, UNISON join NATFHE and AoC on petition, FEnews.co.uk, Thursday 3rd November.

Maynard, J & Smith, V 2004, Practical ways of improving success in modern apprenticeships, Education + Training, vol. 46, no. 5, pp. 253-61.

McAdam, R & Crowe, J 2004, Assessing the business and employee benefits resulting from the implementation of NVQs, Education + Training, vol. 46, no. 3, pp. 138-52.

Messick, S 1989, 'Validity', in Educational Measurement (3rd ed.), Ed R L Linn, Macmillan, New York.

Misko, J 2006, Vocational education and training in Australia, the United Kingdom and Germany, National Centre for Vocational Education Research, Station Arcade, Australia.

Murphy, R, Warmington, P & Macintosh, H 2003, Fair assessment and the discourse of equal opportunities: describing inter-relationships in the context of vocational qualifications, paper presented to the seventh annual conference of the Learning and Skills Research Network, University of Warwick.

Murphy, R, Burke, P, Content, S, Frearson, M, Gillespie, J, Hadfield, M, Rainbow, R, Wallis, J & Wilmut, J 1995, The Reliability of Assessment of NVQs. Report to the National Council for Vocational Qualifications, School of Education, University of Nottingham.

Oates, T 2004, The role of outcomes-based national qualifications in the development of an effective vocational education and training system: the case of England and Wales, Policy Futures in Education, vol. 2, no. 1, pp. 53-71.

Partington, K J 2003, 'Because of x, y and z…' What exactly are x, y and z? Emotional, cognitive and behavioural responses from student teachers in post-compulsory education to teaching students with disabilities, Research in Post-Compulsory Education, vol. 8, no. 3, pp. 407-24.

Pratten, J D & Curtis, S 2002, Attitudes towards training in UK licensed retail: an exploratory case study, International Journal of Hospitality Management, vol. 21, no. 4, pp. 393-403.

Qualifications & Curriculum Authority 2004, The statutory regulation of external qualifications in England, Wales and Northern Ireland, QCA, London.

Queensland Government 2005, Graded assessment in Queensland – a discussion paper, Queensland: Department of Employment and Training.

Raffo, C 2003, Disaffected young people and the work-related curriculum at key stage 4: issues of social capital development and learning as a form of social practice, Journal of Education and Work, vol. 16, no. 1, pp. 69-86.

Smith, J 2005, 'This is our chance to pioneer pathways of progression', Education Guardian, Friday 25th November.

Spielhofer, T & Sims, D 2004[a], Modern apprenticeships: hitting the target? Education + Training, vol. 46, no. 3, pp. 112-18.

Spielhofer, T & Sims, D 2004[b], Modern apprenticeships in the retail sector: stresses, strains and support, Journal of Vocational Education and Training, vol. 56, no. 4, pp. 539-58.

Stanton, G 2004, The organisation of full-time 14-19 provision in the state sector. Working Paper 13, Nuffield Review of 14-19 Education and Training, University of Oxford.

Stasz, C & Wright, S 2004, Emerging policy for vocational learning in England: will it lead to a better system? Learning and Skills Research, Centre London.

Stasz, C, Hayward, G, Oh, S-A & Wright, S 2004, Outcomes and processes in vocational learning: a review of the literature, Learning and Skills Research Centre, London.

Swailes, S & Roodhouse, S 2003, Structural barriers to the take-up of higher level NVQs, Journal of Vocational Education and Training, vol. 55, no. 1, pp. 85-105.

Thomson, P, Mathers, R & Quirk, R 1996, The grade debate, National Centre for Vocational Education Research, Adelaide.

Tolley, H, Greatbatch, D, Bolton, J & Warmington, P 2003, Improving occupational learning: The validity and transferability of National Vocational Qualifications in the workplace. Research Report No. 425, Department for Education and Skills, Nottingham.

Torrance, H & Coultas, J 2004, Do summative assessment and testing have a positive or negative effect on post-16 learners' motivation for learning in the learning and skills sector? A review of the research literature on assessment in post-compulsory education in the UK, Learning and Skills Research Centre, London.

Turner, D 2002, Employability skills development in the United Kingdom, National Centre for Vocational Education Research, Station Arcade, Australia.

Unwin, L, Fuller, A, Turbin, J & Young, M 2003, What determines the impact of vocational qualifications? A literature review. Research Report No 522, Department for Education and Skills, Nottingham.

Whittaker, M 2005, Colleges need to keep an eye on the pupil, The Times Educational Supplement FE Focus, Friday 2nd December.

Whittaker, M 2006[a], Colleges stung by 14 to 16 courses, The Times Educational Supplement FE Focus, Friday 10th February.

Whittaker, M 2006[b], Now the kids are on campus, The Times Educational Supplement, Friday 17th March.

Whittaker, M 2006[c], Wayward pupils a cause for concern, The Times Educational Supplement FE Focus, Friday 31st March.

Williams, M & Bateman, A 2003, Graded assessment in vocational education and training, National Centre for Vocational Education Research, Station Arcade, Australia.

Wolf, A 2002, Same courses, different name: will vocational GCSEs bring credibility to the alphabet soup of qualifications? The Guardian, Tuesday 28th May.

Wolf, A & Jenkins, A 2002, The growth of psychometric testing for selection. Why has test use increased, will growth continue, and what does this mean for education? Centre for Economics of Education, London School of Economics and Political Science, London.

Wolf, A & Silver, R 1986, Work-based learning: Trainee assessment by supervisors, MSC R&D Series No. 33, Sheffield

Wolf, A & Silver, R 1995, Measuring 'broad skills': the prediction of skill retention and transfer over time, Department for Education & Employment Research Series No. 61, Sheffield.

Worth, S 2002, Education and employability: school leavers' attitudes to the prospect of non-standard work, Journal of Education and Work, vol. 15, no. 2, pp. 163-80.

Yaphe, J & Street, S 2003, How do examiners decide? A qualitative study of the process of decision making in the oral component of the MRCGP examination, Medical Education, vol. 37, no. 9, pp. 764-71.