What research could an Awarding Body carry out about NVQs?

Jackie Greatorex

University of Cambridge Local Examinations Syndicate

A paper presented at the British Educational Research Association Conference, University of Cardiff, UK, 7-9 September 2000.

Disclaimer
The opinions expressed in this paper are those of the author and are not to be taken as the opinions of the University of Cambridge Local Examinations Syndicate (UCLES) or any of its subsidiaries.

Note
This research is based on work undertaken by the University of Cambridge Local Examinations Syndicate for Oxford Cambridge and RSA Examinations (OCR).

Contact details
Jackie Greatorex
Research and Evaluation Division, University of Cambridge Local Examinations Syndicate, 1 Hills Road, Cambridge, CB1 2EU.
☎ 01223 553835
FAX: 01223 552700
💻 greatorex.j@ucles.org.uk

Autobiographical details
Jackie Greatorex is a Research Officer at UCLES.

What research could an Awarding Body carry out about NVQs?

Abstract

The National Council for Vocational Qualifications (NCVQ) was instituted in 1986 to reform vocational qualifications. Its role was to rationalise the existing structure of vocational qualifications and to design, monitor and adapt as necessary a new framework of qualifications - National Vocational Qualifications. Awarding Bodies are responsible for the implementation of individual NVQs. They undertake external verification to ensure that candidates are being assessed fairly and consistently across centres. National Vocational Qualifications (NVQs) are audited by Awarding Bodies via a process of internal and external verification, as well as by the QCA (Qualifications and Curriculum Authority), OFSTED (Office for Standards in Education), the TSC (Training Standards Council) and the FEFC (Further Education Funding Council). Within this context and the emerging new focus of Awarding Body self-assessment, the question is what are the most important areas of the delivery of the NVQ on which Awarding Bodies could themselves be focusing their research and evaluation efforts?

KEY WORDS: Awarding Body, NVQ, vocational qualifications.

Introduction

Before National Vocational Qualifications (NVQs) were introduced there was a range of vocational qualifications offered by a range of assessment providers with no coherent national framework and an apprenticeship system which was criticised for being based upon 'time served' rather than competence (Raggatt and Williams, 1999). The National Council for Vocational Qualifications (NCVQ) was instituted in 1986 to reform vocational qualifications. It had:

> *the job of 'rationalising' the existing structure and was directed to ' design', 'monitor' and 'adapt as necessary' a new framework of qualifications - NVQs* (Raggatt and Williams, 1999, 13).

The Manpower Services Commission (MSC) argued that using standards rather than time served would widen access to training and opportunities. The standards were to be based on the knowledge and skills required by the job. The standards would be explicit, agreed, widely accessible, flexible and testable (Raggatt and Williams, 1999). NVQs are assessed against criteria to ensure that candidates have reached the required standards. The criteria are linked to occupational standards. There is no official syllabus for NVQs, although there is an official pedagogy; the awards are meant to reward candidates for taking responsibility for their own learning (Wolf, 1998). To achieve a NVQ a candidate must have met all the criteria specified for the individual qualification.

The assessment of NVQs uses a variety of strategies, portfolio assessment, tests and simulations. NVQs include a variety of assessment decisions which need to be recorded. This led to the dominance of portfolio assessment. The portfolio is a convenient way of evidencing competence and recording judgements (Eraut et al., 1996). Wolf (1998) argues that portfolios are important because they can be more easily verified than other forms of assessment e.g. ephemeral performance.

The NVQ quality assurance system involves a large network of individuals and organisations

whose roles are described below.

Who does what in the current system?

The roles of organisations in the current system are as follows:-

- *Qualifications and Curriculum Authority* (QCA) ensures that NVQs meet particular criteria and are broadly comparable across different sectors. QCA accredits (formally recognises) proposals for NVQ awards developed by National Training Organisations and Awarding Bodies and quality assures and audits the quality assurance activities of the Awarding Bodies (QCA, 2000c).
- *Standards Setting Bodies*, normally National Training Organisations (NTOs), identify, define and update employment-based standards of competence for agreed occupations. NVQs are based on these standards and criteria established by QCA (QCA, 2000c).
- *Awarding Bodies* (ABs) have a dual role. With standards setting bodies, they jointly develop NVQs and are also responsible for the implementation of individual NVQs. The ABs approve centres who wish to offer NVQ assessment services. Awarding Bodies monitor the quality assurance assessment processes and award NVQs and unit certificates. They undertake external verification to ensure that candidates are being assessed fairly and consistently across centres (Konrad, 1998) against the Common Accord Criteria and they have nationally agreed procedures for the verification system (QCA, 2000c). The Common Accord contains detailed quality assurance requirements for implementing NVQs.
- *Centres* appoint their own assessors and Internal Verifiers for each qualification. The latter verify that assessment decisions are valid.
- *Assessors* are competent according to specified occupational standards and should hold the Units D32 (Assess candidate performance) and D33 (Assess candidates using different sources of evidence). These units are part of the NVQ framework. An assessor cannot 'sign off' units for a candidate unless the assessor has the appropriate 'D' units. They judge candidates' performance and knowledge evidence against the NVQ standards, provide constructive feedback to their candidates and sign off candidates' completion of their NVQ. They maintain the level of their professional competence (Konrad, 1998).
- *Internal Verifiers* (IVs) are competent against occupational standards and are qualified as an Internal Verifier i.e. they hold D34 (Internally verify the assessment process). IVs ensure the quality of assessment judgements and processes within an approved centre. They select and train assessors and monitor assessment by sampling candidate evidence of competence and maintaining and developing the quality of assessment and verification documentation. They authorise requests for the award of NVQs and ensure equality of access to assessment for all candidates (Konrad, 1998).
- *External Verifiers* (EVs) are appointed by Awarding Bodies. EVs are required to hold the D35 unit (Externally verify the assessment process). EVs ensure the reliability and validity of the assessment and internal verification process, across NVQ centres, in accordance with the Common Accord criteria. EVs also provide guidance and support to centres. The process of external verification involves sampling assessment and verification practice, providing feedback and an EV report to the centre and the appropriate Awarding Body. The verification chain, specifications and evidence of achievement, are viewed by some, for example, Eraut et al. (1996) to be paper dominated, and they conclude that ultimately the cornerstones of the NVQ, i.e. specification and verification, do not necessarily guarantee standards and fairness.

Awarding Body Research

Examination Boards and Awarding Bodies have undertaken research for a number of reasons. For example, Stephenson et al. (1999) were commissioned to undertake research by RSA partly because the research itself was inherently useful and partly because it was a good way of promoting RSA. In a similar way research staff at UCLES (University of Cambridge Local Examinations Syndicate) undertake research relevant to the organisation's work and are encouraged to publish research work where possible to gain publicity and promote UCLES and its subsidiaries as a reputable organisation which bases its practice upon research evidence. For example, the organisation can claim that its products and services are underpinned by research. Hammersley (2000) described different modes of research. Mode 1 research is controlled and co-ordinated by the researchers. It is the pursuit of knowledge for its own sake and there are parallels with the market but it is not market driven. Mode 2 research is focused on solving problems, undertaken by teams, transdisciplinary, nonhierarchical and product orientated. It is often undertaken so that organisations can claim that their product is based on research (Hammersley, 2000). Research in Examination Boards and Awarding Bodies has features of both mode 1 and mode 2 research.

Using UCLES as an example of an Examination Board the research effort has tended to relate to two main themes:-

1) equity, comparability, fairness (e.g. reliability and validity) or social justice;
2) the quality of assessment and its associated processes.

Usually this work has tended to focus upon school examinations. As described below, however, with the advent of reorganisation and the formation of OCR the focus of the UCLES' research effort is likely to change to some extent.

There are of course limitations to 'in-house' research and evaluation. For example, the research funded by an organisation tends to be biased towards the interests of that organisation. Bates (1995) considers some of these issues. She notes that much of the evaluation of NVQs was sponsored by the National Council for Vocational Qualifications (NCVQ) (the organisation instituted to reform vocational qualifications) and the Department of Employment. Bates (1995) adds that this was a considerable investment. She argues that the research undertaken by such large organisations is likely to be uncritical and focus on managing change rather than undertaking academic enquiry and that they will:

> *Define reality rather than question its social construction* (Bates, 1995, 8).

However Bates (1995) adds that:

> *…despite these constraints, there are signs that the sheer weight of findings and growing experience have begun to lead to some questioning of fundamental precepts* (Bates, 1995, 8).

This suggests that despite the obvious limitations of ‘in-house’ evaluation and research the assumptions behind an assessment system can be questioned and the evidence can be used to instigate radical improvements.

When Awarding Bodies offer NVQs they are audited through a process of external verification. NVQs are also audited by QCA (Qualifications and Curriculum Authority), OFSTED (Office for Standards in Education), the TSC (Training Standards Council) and the FEFC (Further Education Funding Council). In recent years there has been considerable reorganisation of examination and

assessment bodies in the UK. Part of this reorganisation has involved mergers of bodies offering academic and vocational qualifications. In 1998 a new Awarding Body Oxford Cambridge and RSA Examinations, was formed, bringing together vocational and academic organisations. Given the formation of OCR its research scope has expanded to include both 'academic' and 'vocational' qualifications. Within this context the question addressed in this paper is what research and evaluation could an Awarding Body conduct about NVQs?

NVQ Research

This section of the paper deals with the research that has been undertaken about NVQs which is related to issues that could be researched by Awarding Bodies.

NVQ take-up

One possible area of research would be to look at NVQ take-up. There are currently three main types of qualification available to 16-19 year olds - Advanced GNVQs (to be called Vocational A levels), A levels and NVQs. The first two are set to change from September 2000. For further details see QCA (2000a) and Hodgson and Spours (2000).

NVQs are intended for candidates who have left full time education (Smithers, 1994). Robinson (1996) reports that over one-third of those working towards a NVQ in the Spring of 1994 were aged 20-24. He adds that in 1992-1993, 1993-1994 and 1994-1995, 34-40% of NVQs were awarded to 16-19 year olds. 428292 NVQs were awarded in the year 30 September 1998 to 30 September 1999 (QCA, 2000b). Unfortunately QCA (2000b) does not indicate how many of these NVQs are awarded to young people but these figures do give an indication of the popularity of NVQs. QCA (2000c) report on their website (http://www.qca.org.uk/shframes.htm) that:

> *The total number of NVQ certificates awarded to 31 December 1999 was over 2.7 million (an increase of over 19% on the total awarded to 31 December 1998).*
>
> *The number of NVQ certificates awarded in the 12 months to 31 December 1999 fell to 430,000, a decrease of 3% on the 12 months ended December 1998 (442,000).*

It might be this decrease in NVQ take-up which has encouraged the DfEE to support the idea of the Technical Certificate. Related Vocational Qualifications were mentioned in the Beaumont Report (1995). Subsequently they have become known as Technical Certificates. They would involve more external assessment of knowledge. Arguably external assessment increases the rigour of the qualification. The skills involved in these occupations would be assessed using other methods. At the moment discussions are taking place to explore ways that Technical Certificates might be implemented. However the Technical Certificate still remains at the ideas stage.

But what makes people choose to undertake NVQs? Eraut and Steadman (1998) interviewed staff at ABs, accredited centres and qualification authorities and surveyed 215 level 5 Management candidates to evaluate higher level Scottish Vocational Qualifications (SVQs) and NVQs. Some of their findings identified the reasons why candidates chose to do a NVQ in Management. 49% of the candidates claimed that the first initiative in working for a NVQ had come from their employer and 44% said that it was their own initiative. They also found that the employer support for NVQs varied with sectors. Only 6% of respondents thought that their employer valued NVQs more highly than academic qualifications. It appears that employers were not asked about these issues. Respondents were asked to rate 11 factors in influencing their decision

to work towards a NVQ. The four highest rated factors were 'career plans', 'financial support', 'the NVQ emphasis on competence' and 'discussion with the training provider'. The low rated factors were 'result of formal appraisal' and 'employer's willingness to grant time off'. The other factors were 'advice from a professional institute', 'employer requires it for your grade', 'example/advice of colleagues', 'increased chances of getting another job' and 'increased chances of promotion'. There were variations in these ratings within different employment sectors. Although this identifies some of the reasons for undertaking a NVQ it does not suggest what can be done to increase the take-up of NVQs as this was not part of the research brief.

Stephenson et al. (1999) report that people who achieved NVQs experienced enhanced self-esteem, personal qualities and skills. They also found that candidates were not generally self-motivated in terms of beginning and finishing an award. Candidates were encouraged by those with executive responsibility to begin the award. This was different to Eraut and Steadman's (1998) finding that over 40% of candidates had initiated working for a NVQ. The difference between these findings might be explained by the higher level management candidates surveyed by Eraut and Steadman (1998) being expected to take more responsibility for their own professional development than the candidates whom Stephenson et al. (1999) interviewed in their study, which incorporated a variety of sectors and NVQ levels. Stephenson et al. (1999) add that some candidates dropped out when this pressure from executives eased. Those who were internally motivated were kept going by personality/personal interest, and sometimes this was to prove a point.

QCA (1999b) report that at NVQ user forums it was suggested that smaller packages of NVQs would increase NVQ take-up. However some of the NVQ users considered such packages unsuitable for young people beginning their career as they would need to demonstrate the full range of NVQ competence. These qualifications would be more suitable for people who had a great deal of work experience and wanted to develop their skills further in one area.

The DfEE are asking for bids to undertake research for their 2000-2001 research programme. One of the projects in the programme would include statistical and qualitative research to identify why the take-up of NVQs has reduced amongst younger candidates and how to address this. The research would complement work which is already being undertaken about why the number of older candidates for NVQs has reduced (DfEE, 2000).

The research discussed above suggests that research about why those with executive responsibility support candidates taking NVQs and what an AB could do to encourage this would be as worthwhile as research about why candidates take a NVQ. It is important that an AB is aware of the reasons for candidates undertaking and not undertaking NVQs and knows what strategies it could use to increase the number of candidates taking appropriate awards. Of course NVQs are developed, administered and awarded by several organisations so if the problem of low take-up lies in the awards then this is difficult for one AB alone to rectify. Any improvement would need to be instigated in partnership with QCA and the appropriate NTO and other ABs. If the problem lies with the marketing of the NVQ or another issue which is related to AB responsibilities, then the research might be used to improve AB practice.

The External Verification Process

An important area in the NVQ system is the EV system. The EV acts as the point of communication between the centres and the appropriate AB. Eraut et al. (1996, 65) discussed the results of their survey:

> *The results...must be of very considerable concern, in particular our finding that 38% of assessors feel that many candidates pass who shouldn't....Most astonishing are the higher figures of 41% for internal verifiers and 48% for external verifiers, suggesting that, the greater their knowledge and experience of the NVQ assessment, the more likely it is that assessors feel that wrong decisions are being made.*

They also describe EVs as those most involved in implementing the NVQ system.

QCA (1999b) report that they are developing a Code of Practice for external verification for which there is considerable support. The Code is being developed in consultation with ABs. At NVQ user forums a number of issues were raised by users about the experience of being externally verified. The users considered inconsistency in EV practice as a serious problem. For example, the standards of different EVs were thought to vary between ABs, and within ABs, for the same qualification. There were also concerns about the different recording systems that the EVs (and ABs) required. Centres were interested in having an outline of what they should expect from an EV. It is hoped that the new Code of Practice and standardising the guidance from ABs to EVs within each sector will tackle some of these problems. QCA (1999c, 1) state that:

> *The Awarding Bodies used external verification to support assessment consistency, however most needed to introduce formal national reviews in this area to inform future arrangements. Although many awarding bodies had improved their arrangements for external verification, further work was still required to ensure that the performance of external verifiers was effectively evaluated. This would have a particular impact in improving the quality of centre approval and monitoring and the consistency of the award.*

They added that evidence from centres showed that 81% thought ABs provided them with acceptable assessment guidance. Whilst ABs monitored assessment at the centre level through EVs QCA (1999c, 2) report that ABs:

> *did not have the ability to monitor and review assessment issues centrally to indicate where improvements were necessary to the NVQ and / or its assessment guidance.*

QCA (1999c) also found that the majority of the ABs did not comprehensively monitor their EVs. In the 62% of centres visited by QCA external verification failed to meet one or more of the requirements in the Common Accord.

The other side of the coin is the experience of being an EV. The NVQ user forums also mentioned:

> *the problems caused by high turnover of external verifiers* (QCA, 1999b, 18).

One possible way of overcoming these problems that was suggested was making external verification into a full time profession rather than a part-time occupation. However this would cause difficulties in maintaining the occupational competence of EVs.

The studies considered above suggest that the external verification process needs further scrutiny. QCA (1999c) do not make clear to what extent the issues raised by the monitoring apply to each AB. Bearing this in mind an AB could approach centres to investigate the experience of external verification through the eyes of the centre. The research outcomes could be used to improve the AB's products and services. It would be appropriate to try to undertake this work as a national research project so that the evidence could be centrally collated and analysed. The literature that

has been reviewed also suggests that research into the experience of being an EV and their views about the occupation might provide evidence which could be used to improve Awarding Body external verification.

External Verifier Training and Judgement

This section concerns external verification but first relevant research about assessor judgement is reviewed as there is little research about verifier judgement. An important issue in assessment is assessor judgement. In the context of school examinations assessor judgement was researched by Examination Boards, for example, Baird (1999) and Scharaschkin and Baird (2000). Eraut et al. (1996) and Murphy et al. (1995) also undertook work in this area of assessor judgement. They focused upon reliability and what affected the consistency of assessor judgement. Baird (1999) and Scharaschkin and Baird (2000) also considered factors which might affect assessor judgement. Eraut et al. (1996) also considered amongst other issues the criteria that some assessors used to make judgements besides the stated NVQ criteria.

Wolf (1995) explains that standards are carried in examples of work rather than in assessment criteria because if assessment criteria are separated from students' work they can be interpreted as appropriate for many different levels of achievement. Of course different users read exemplars differently. There is little research about the efficacy of exemplars in creating common understandings of standards (Wolf, 1995). Black et al. (1989) report that in Scottish National Certificate modules there was a communication module. The assessors had found it difficult to interpret the criteria so they founded a network where standards were discussed. This led to a common understanding of the criteria. This in turn led to consistency of judgement. However it is logistically difficult for NVQ assessors to have this shared assessment experience which would help them to internalise the standards. This is why the internal and external verification process is important.

Konrad (1998b) describes the notion of cognitive apprenticeship (sharing the culture of real practice using knowledge as a tool for students to use in a collaborative social interaction). He states that:

> *I believe that it would provide a more coherent and effective approach to training and development than the limited approach based on the current "D units"* (Konrad, 1998b, 4).

He also argues that an approach of collaborative working, i.e. team work rather than hierarchical organisation, would improve the system. Konrad (1998b, 5) argues that:

> *a group of assessors and verifiers may be regarded as a community of practice if they share common values, purposes and practices that have a direct practical effect on their professional and personal actions.*

His paper is mostly about IVs, but the debates might also be applied to EVs. Eraut and Steadman (1998) agree with Konrad (1998b). They conclude that the training of assessors and verifiers (D units) is not focused sufficiently on building assessment communities which will result in consistent comparable judgements. It appears that there are moves towards developing this kind of community. For example, QCA (1999b) report that there was support amongst NVQ users for a professional body for EVs which might help with effective networking. Judging by Wolf's (1995) argument this networking could be a key to consistency of judgement.

Besides holding the appropriate 'D units', EVs receive training before they undertake verification. In meetings they are given evidence similar to that which they would be likely to verify. If there are any differences in opinions between the verification judgements these are addressed. This process is undertaken to ensure that the EVs interpret the NVQ standards in a similar way. The training is evaluated. Although EVs are given this training QCA (1999c) reported that most ABs do not provide their EVs with comprehensive professional development. There is little research undertaken about the similarity of EV judgements after training and how the training translates into practice. QCA (1999b) report that a new EV professional body is emerging. Research such as that mentioned above would help improve EV training and might provide useful information for the EV professional body.

External Verifiers and Quality Assurance

RSA commissioned Stephenson et al. (1999) to undertake some research about the use of NVQs as a means to develop corporate capability and the contribution of NVQs to the learning milieu of the work place. Stephenson et al. (1999, 14) report amongst other things that their empirical work showed that:

> *There was almost universal preference for internal assessors who knew what the job entailed and who as a consequence had some credibility in the eyes of the candidates.*

In one of the companies which was part of the study it was found that there was a preference for assessors who had not only done the job but who were still called upon to do it. The internal assessor was seen to have an important role in the development stage of candidates. The internal assessor provided opportunities for the candidates to talk about learning, achievements and further progression which stimulated motivation. Two of the companies had unfortunate experiences with external assessors who did not seem to know what was required by the job. These assessors lost the confidence of candidates. Stephenson et al. (1999) do not give an account of the companies' reactions to IVs and EVs as this was not part of their research brief. However it seems possible that companies might hold the same views of EVs and IVs as they did of external assessors.

QCA (1999b) report that there are concerns about the occupational competence of EVs. The NTOs do, however, specify occupational criteria for the appointment of EVs. There is also a requirement for EVs to maintain their occupational and professional development. In addition, others consider that EVs should hold an Assessor Award (D32 and D33), and an IV Award (D34) besides the mandatory D35 (EV Award). At the NVQ user forums QCA (1999b, 18) report that:

> *The possibility of increasing the number of full time 'professional' external verifiers was raised, although it was recognised that continuing professional development would be needed to keep occupational competence up to date.*

This direction would, however, mean that EVs would have less opportunity to maintain experience through work. This scenario leaves room to investigate the factors which affect the effectiveness and professional enhancement of the EV.

EV Reports

When EVs check the quality of assessment and processes in a centre they report this to the AB using an EV Report form. This report contains information which identifies whether the criteria

in the Common Accord have been met. It was designed to enhance the quality, cost effectiveness and coherence of the assessment and verification of NVQs. In the report the EV also records the outcomes of the verification process. When the AB receives these forms the contents are considered and any necessary actions taken to improve practice at the centre level. QCA (1999b) found that there were requests for ABs to incorporate centre feedback on EVs' forms as part of the monitoring process. Other than this the issue of EVs' reports appears to have been neglected in the literature, probably because this is mostly of interest to centres and ABs. The qualitative comments that are made in the reports could be analysed to identify the issues raised. This could be undertaken in one sector or a range of sectors. The information would be used to evaluate the efficiency and effectiveness of the Reporting Form. It might also provide insights about best practice which would be of value to centres and could be used as part of the guidance that ABs give to centres.

Transfer

Credit Accumulation and Transfer systems (CATS) have been used in Higher Education for some time. In these systems units of qualifications or whole qualifications are given a credit tariff and credits can be accumulated to gain a whole qualification. NVQ units are of different sizes and therefore it would be difficult to introduce a credit rating system. The issue of transfer is related to credits. If whole qualifications or units are credit rated then a candidate can take credits from one educational institution to another and still gain an award. The credits might be transferred from, say, a BSc in Psychology to one in Social Sciences provided that the outcomes of the two programmes at the point of credit transfer were sufficiently similar. This principle also applies to other qualifications. QCA (2000c) state on their website that:

> *The primary purpose of the NVQ framework is to create a coherent classification for NVQs and to facilitate transfer and progression, both within areas of competence and between them… Further refinements to the system are being made as the qualifications are developed and routes for progression/transfer are identified.* (author's emphasis).

The notion of credit transfer is also linked to the concept of transferring learning from one context to another or learning and skills which are relevant to more than one context. Key Skills previously known as Core Skills are regarded as relevant, and hence perhaps transferable, across a wide range of contexts.

In the early 1990s NCVQ were involved in a pilot programme to accredit the Core Skills of 'communication, problem solving, personal skills, numeracy, information technology and modern language competence'. In 1993 the Employment Department's Methods Strategy Unit undertook a short but detailed study about transferable skills. They aimed to identify best practice for enhancing transferable skills as implied by the NCVQ Core Skills, to identify current practice and know-how in all relevant settings including schools, colleges, employment organisations and other training providers and to review, identify and illustrate best practice. The Employment Department established a network of experts, reviewed literature, surveyed good practice, ran a consultative workshop and visited practitioners. The survey was based upon examples which claimed to be good practice. 41 pro-formas were completed by experts. The results suggested that within each education sector e.g. FE and Core Skill area the feedback from the survey was similar. They found that there was much confusion and disagreement about Core Skills and transfer issues. All of the NCVQ Core Skills had transfer value. But having a 'transferable' or 'core' skill does not mean that the skill can be used in novel or demanding circumstances. Rather it can be used in a fresh context. From their literature review they found that there was an increasing consensus on the learning environments likely to enhance transferable skills. They

added that there was a gulf between desirable and actual practice. This was due to a number of factors which included organisational constraints and poor awareness of the value of Core Skills. They offered some guidance for good practice (Employment Department, 1994).

Anderson and Marshall (1997) argue that changes in technology and the economic environment mean that people require more non-occupational specific skills. They query whether even technical skills can be considered as occupationally specific. Employers attach importance to the Core Skills (now Key Skills) which are incorporated in some NVQ units, although some employers have their own lists of the Core Skills which they want. For example, British Telecom (BT) expects all employees to be flexible and have the ability to adapt and cope with uncertainty. Employers complain that some groups, for example, some graduates do not have the personal skills required to contribute quickly and effectively in the workplace and find difficulty in writing sentences, presenting arguments and spelling. The Employment Department's publication *Thinking and Learning at Work* makes the underlying assumption that people can be trained to develop Key Skills. However, Anderson and Marshall (1997) raise the question of whether Key Skills can be taught. They conclude that traditional academic approaches to learning and assessment are inhibiting factors and that some skills are more or less trainable. They also added that transfer is limited, which was also acknowledged by the Employment Department (1994).

In general, NVQs are occupationally specialised, for example, 'Marine Operations (Harbour-based)' level 2 or 'Coke Making' levels 2 and 3. In order to make NVQs more accessible a core and options model was introduced. That is, there are some units which are common to a number of qualifications and then if certain optional units are chosen the candidate can be awarded a particular qualification. This kind of rationalisation is attractive as there are many NVQs, some with few candidates. For example, 'Quality Management' at level 4 has been awarded to 27 candidates and 'Piloting Transport Aircraft' level 4 has been awarded to 21 candidates, and in the 'Tending animals, plants and land' sector, 50% of the qualifications were each awarded to less than 100 candidates. These figures were for all NVQs, both current and expired as at 30 September 1999 (QCA 2000b). Of course it would not be appropriate for qualifications to be designed with common units if these units were not relevant to each award.

QCA (1999b) explored these issues through NVQ user forums. There was strong support for the core and options structure as it offered a high degree of flexibility. This flexibility would need to be balanced with the quality assurance system - it would be difficult to find assessors and verifiers who had the occupational competence to verify some or all of the whole awards which would emerge. There were various views on how large the 'core' should be. The general agreement was that NTOs should specify rules for combining core and option units and determine the size and content of the core (QCA, 1999b). QCA (1999b) also reported support for a greater focus on generic skills within NVQ core units. These generic skills would be transferable across a range of occupational areas. The current Key Skills units were seen as a good example and it was suggested that health and safety units could be included in the core. There was also support for the notion of importing units from already existing NVQs into new NVQs as this facilitated transfer. The NVQ user forums mentioned that the other key function of NVQ core units was to set the transferable national standard for the award (QCA, 1999b).

This reopens the debate about transfer and how generic some skills can be whilst remaining relevant to a series of qualifications and specialist occupations. The feasibility of using the core plus options model for other qualifications could be investigated. Any research in this area would need to take into account the views of relevant NTOs concerned with the development of NVQs.

Capability

In social cognitive theory behaviour, cognitive and other personal factors all affect one another. Bandura (1988) explains how they can be used to improve organisational functioning. His three strands of discussion are:-

- guided mastery modelling;
- strengthening peoples' beliefs in their capabilities;
- enhancing self-motivation through goal systems.

Guided mastery modelling is modelling appropriate skills. More specifically this is people receiving guided practice under simulated conditions and support to apply their skills in the workplace. Bandura (1988) explains that this is an optimal way of supporting people as they reach competence.

Self-efficacy is the extent to which people believe in their capabilities to exercise control over events to accomplish goals. Self-efficacy affects how well people perform by affecting their motivation and problem solving efforts. In short self-efficacy directly affects capability. Everyone will experience failure and setbacks which can lead to self-doubt. Therefore it is resilient self-belief that counts. Bandura and Wood (1989) found that if managers believed that organisations were controllable then they maintained a robust sense of managerial efficacy and adopted high organisational goals even when these goals did not come easily. Managers who did not believe that organisations were controllable soon gave up trying to achieve goals when they were not easily reached. They also found that initially mangers will judge their capabilities by their past performances but with further experience beliefs about capabilities became a more powerful determinant of their organisational attainments. Such studies reveal that organisational achievement is strongly affected by managers' beliefs about their capabilities rather than the abilities of the employees who make up the organisation.

Social cognitive theory illustrates that people exercise self-directedness and adopt internal standards, they monitor their behaviour and use incentives to motivate themselves to accomplish their goals. Goals are highly effective motivators if they match peoples' perceived capabilities. Goals affect motivation through self-evaluation, how people react to their performances and people's beliefs in their capabilities.

In summary social cognitive theory provides guidelines on how to develop people's self regulatory capabilities and a resilient sense of self-efficacy which enables them to enhance their personal accomplishments (Bandura, 1988). This is a crucial point in attempts to increase corporate capability - the sum total of the individual capabilities of the employees in an organisation. Increasing the resilience of employees' self-efficacy will enable them to enhance their personal accomplishment which will increase the achievements of the organisation (Bandura, 1988).

Organisational Learning and Capability

DiBella et al. (1996) found that there was a variety of different ways for describing organisational learning capability and understanding learning styles. It was also reported that organisational learning might be increased by building on existing capabilities or developing new ones. The first involves improving current capabilities and the second involves a change in culture. Organisations can improve their learning capability through either approach.

NVQs and capability

RSA commissioned two research projects through Higher Education for Capability:-

- *The Contribution of National Vocational Qualifications (NVQs) to the learning milieu of the workplace* by John Stephenson, Ralph Williams and Peter Critten (1999);
- *The Use of NVQs as a Means to Develop Corporate Capability* by Ralph Williams, Lynne Cunningham and John Stephenson (1997).

Williams et al. (1997) identified a key feature of organisational capability as a 'healthy learning milieu'. Both corporate capability and learning are linked to the notion of a 'learning organisation'. They raised questions about the role of NVQs in the development of corporate capability. This work was based on three case studies. Stephenson et al. (1999) provided a checklist of organisational features which indicated whether NVQs have been successfully implemented in a company. The report was based on 10 organisations and covered the experiences of 200 people. Stephenson et al. (1999) identified 8 features of a healthy learning milieu. Two features were commonly associated with NVQ activities: *learning leads to personal growth* and *learning interacts with work.* For the other features - *learning is shared, learning is prioritised, learning is own responsibility, learning is for all, learning pays* and *learning is continuous* - there were varying degrees of overlap between NVQ activities and a healthy learning milieu. NVQs were more likely to be associated with organisation driven learning activities than learner driven activities. This is an interesting point as Wolf (1998) argued that the unofficial pedagogy of NVQs is that the learner is responsible for their learning. But Stephenson et al. (1999) seemed to imply that the organisation rather than the learner was driving the learning.

The work commissioned by RSA has focused upon the benefits of NVQs for corporations and identified that there are significant benefits to the corporation and the individual e.g. certification, learning, promotion and personal growth. Jackson and Jordan (1999) considered who benefits from skills training. Increasing one's qualifications will only subtly enhance one's position in the labour market. So whilst the responsibility for learning in qualifications like NVQs is ideally shouldered by the candidate it is both the employer and candidate who experience gains. Additionally Inge Bates (1999) found that work is increasingly overlapping with people's private lives. The RSA work hints that NVQs can infringe on people's personal lives. There is room for further investigation here – to what extent do work and NVQs overlap with people's personal lives? Do the costs of this overlap justify the benefits of undertaking the qualification? What benefits do candidates (rather than corporations) gain from undertaking NVQs? (The research commissioned by RSA does provide some answers, e.g. personal growth). How can the learner rather than the organisation be encouraged to take responsibility for their own learning by driving the learning and would this be desirable?

Reliability and Validity

As explained earlier Examination Boards have a history of evaluating the measurement characteristics of school examinations. These methods are unlikely to be directly applicable to NVQs which are entirely criterion referenced whereas for A level and GCSEs etc. statistics and criteria are used to determine grade boundaries. Eraut et al. (1996) pointed out that there is a lack of literature in the area of the measurement characteristics of NVQs.

Eraut and Steadman (1998) define the external validity of a competence based award as the match between the competence measured by the award and the performance of people deemed competent in the work place. One of the questions in their survey was *What competencies do you have that fall outside the content of your portfolio?* Only 22% responded *none - portfolio*

coverage is good. The other 78% said that they had other skills e.g. knowledge specific to their business. This is likely to occur with a generic award. 72% of respondents said that their portfolio gave a valid picture of their competence in the areas covered by the NVQ and only 6% reported that it was not valid. They also report that pressures of convenience, speed and probability of success might be more important to candidates and assessors than validity when deciding on an assessment strategy.

The reliability of NVQs depends on the consistency of assessor judgements. The Beaumont Report suggested that there was no major cause for concern about the consistency of judgements. They claimed that people were concerned about reliability because they did not believe that assessors could be objective. No one can be entirely objective but as qualifications are a gateway to jobs in a meritocracy it is important that they are seen to be fair which can mean as objective as possible. Beaumont (1995) said that *The Reliability of Assessment of NVQs* project (undertaken by Murphy et al. [1995]) found that:-

- *there was substantial agreement across assessor decisions – most disagreements were concerned with sufficiency of evidence;*
- *where a number of assessors observed performance, there were only marginal differences in overall judgements made* (Beaumont, 1995, 38).

Eraut et al. (1996) contradicted the first point, saying that there were caveats which need to be considered when interpreting the results of Murphy et al. (1995). The caveats affected whether the agreement could be considered to be 'substantial' (more details are given later). Eraut et al. (1996, 66) investigated the assessment of NVQs. They concluded that:

> *NVQs as currently awarded cannot be said to indicate achievement of nationally consistent explicit standards of competence; and the claim that the assessment of NVQs is valid and reliable cannot be sustained.*

An example of the reliability of NVQs being compromised is assessors taking a wide range of unofficial factors like attitude into account when making assessment judgements. Eraut et al.'s (1996) findings are very important as Wolf (1998) says that they are the results of the only major independent study of assessment decisions in NVQs. Wolf (1998) argues that there is a lack of scrutiny of NVQs because other qualifications like GNVQs are more important in the process of selecting and sorting young people for higher education, and there are difficulties involved in observing NVQ assessment. Despite Wolf's view NVQs are heavily audited by AB external verification, QCA, OFSTED, TSC and FEFC. The consistency of judgements and comparability of standards should have been operationalised through the external verification system. Senker (1996) claims that EVs generally checked that the correct procedures had been followed, rather than checking standards of candidates' performance. This claim was made some time ago and the situation might now be different.

If ABs are to investigate NVQ reliability they would need to find an appropriate method. Murphy et al. (1995) undertook the first reliability study of NVQs (Wolf, 1998). They noted that general theories of reliability and validity do not apply to NVQs as they are criterion referenced, using a competent and not yet competent system. They suggest that it is difficult to compare the reliability and validity of the assessment system with other systems because it is such a different model. Like Eraut et al. (1996) they noted that judgements are affected by a range of contextual factors. The study involved independent assessment of portfolio evidence and assessors discussing the problems that occurred. Participating assessors were asked to bring portfolios which were 'not yet competent', 'just competent' and 'clearly competent'. The assessors made

independent judgements about the portfolios. These were recorded as 'competent', 'competent but uncertain', 'insufficient evidence', 'unsure' and 'not competent'. They also jointly reviewed their decisions and recorded changes in their views. After discussion assessors did not generally change their minds. When Eraut et al. (1996) reviewed Murphy et al.'s (1995) findings they said that in their experience of NVQ assessment 'competent but uncertain' should be classed as 'competent'. Using this definition of 'competent' and 'not competent':

> *of the 35 units assessed by 3 to 7 scrutineers, 16 included a decision of both 'competent' (C, C?) and a decision of 'not competent' (NC)* (Eraut et al. 1996, 11).

What Eraut et al. (1996) do not make clear about Murphy et al.'s work is that they used a few units in a number of different sectors. There were 35 units in total. Each scrutineer assessed a small number of units in their sector of expertise. For example, 4 Vehicle Maintenance scruntineers assessed a maximum of 7 units but one scrutineer only assessed 4 units (Murphy et al.1995). Eraut et al. (1996) argue that if the assessors were considering a large ability range this number of disagreements would be more significant or problematic than if a small ability range were considered. But it is difficult to establish the width of the ability range in Murphy et al.'s (1995) study. Observations of tasks gave higher levels of agreement. To conduct their investigation Eraut et al. (1996) visited 6 contrasting assessment sites where they undertook observation, interviews (with 7 assessors, 17 centre managers and a number of trainees) and gathered documentary evidence. They also received questionnaire responses from 1233 assessors. They concluded that the claim that the assessment of NVQs is valid and reliable cannot be sustained.

QCA's (1999c) comparability studies of the similarity between local assessors' and scrutineers' (occupational experts) judgements involve 229 centres across four sectors - Engineering, Catering and Hospitality, Information Technology, Sports and Recreation. The scrutineers' judgement is considered to be the true score and the decisions of the assessors are the observed scores. Error is the difference between the true and observed scores and the reliability of the assessment is the degree of variance of the errors. The focus is upon summative assessment. All scrutineers collect data in a standard format. It is recognised that the evidence viewed by the scrutineer is a snapshot of what is seen by the assessor. Therefore the assessor and scrutineer are making slightly different types of decisions. The results for the different sectors identify different strengths and weaknesses.

This review has illustrated the methods that were used to evaluate the reliability and validity of NVQs. There are parallels between the Murphy et al. (1995) and QCA (1999a) comparability work – comparing two sets of decisions. Murphy et al.'s (1995) work also involved a discussion session to clear up any differences in judgement. Discussing work to clarify the standard is good practice as it increases agreement between judges (Wolf, 1995). The visits to centres, e.g. QCA (1999a) and Eraut et al. (1996), and / or a survey e.g. Eraut et al. 1996 would add information about the issues which might improve / compromise validity and reliability. Wolf (1995) employed a ranking exercise. At a superficial level this does not fit with the competent / not yet competent dichotomy. However once the ranking was finished the participants could then consider the point in the ranking order at which candidates could be described as 'competent and 'not yet competent'. There might also be some candidates for whom there is uncertainty between assessors about their competence.

An area of potential research for ABs is the validity and reliability of NVQs and methods of evaluating these qualities. Any future AB research about the validity of NVQs should consider the findings of some research in progress commissioned by the DfEE to develop tools to assess

the validity and transferability of NVQs in the workplace and the extent to which employees who possess these qualifications can perform adequately. It is suggested that investigations about the measurement characteristics of NVQs might involve the comparison of assessors' and scrutineers' judgements and visits to centres / surveys for further information.

Conclusion

A series of possibilities for AB research about NVQs are presented. Of course, these suggestions need to be prioritised and methods for undertaking the research considered. Additionally these suggestions will need to be considered in the light of ongoing discussions about how NVQs might be improved. For example, QCA are considering how a Technical Certificate might be implemented. The most urgent research is determined by the requirements of external monitoring, for example, the QCA monitoring. This raises issues about external verification, particularly EV training. Given the QCA monitoring results and the vital role of EVs in the reliability and validity of NVQs, research concerning EV training could be a priority. Whilst research is conducted in one area about EVs, if resources allow, an in depth research programme about external verification could be devised covering the other issues which have been mentioned e.g. the consistency of external verification as perceived by centres and the EV reporting process.

The issue of NVQ take-up is very important to Awarding Bodies. Given the prominence of this issue for ABs, it might come next on an AB research agenda. It is vital that they offer quality products.

In assessment, measurement or fairness issues are a priority. Research into NVQ measurement characteristics should be ranked high in order of importance. It would complement the internal evaluations which are currently undertaken in relation to GCSEs and A levels.

RSA funded research about capability. OCR have now shown an interest in this area of research. A fourth priority might be to build on the research that has already been done. Research of least priority would probably be about transfer. Research into transfer issues would need to be conducted in consultation with NTOs. Of course the NVQ research suggestions would need to be prioritised in relation to other AB research interests. Such a discussion is beyond the scope of this paper.

To identify research priorities for OCR colleagues were consulted and a literature review was undertaken. It is beyond the scope of this paper to present a detailed review of the history of the main issues that have concerned NVQ developers and researchers. However some of the current issues and literature and how they relate to the research that ABs could undertake are given above. Undoubtedly other delegates will be able to identify further projects and associated literature. Any recommendations on these issues are warmly welcomed.

References

Anderson, A. and Marshall, V. (1997) *Core versus Occupation-specific skills.* Department for Education and Employment. www.the-stationery-office.co.uk/document/dfee/resbrief/brief12.htm.

Baird, J. (1999) *Are examination standards all in the head? Experiments with examiners' judgments of standards in A level examinations*. A paper presented at the British Psychological Society London Conference, Institute of Education, London, December 1999.

Bandura, A. (1988) Organisational Applications of Social Cognitive Theory. *Australian Journal of Management*, 13, 2, 275- 302.

Bandura, A. and Wood, R. E. (1989) Effect of perceived controllability and performance standards on self-regulation of complex decision making. *Journal of Personality and Social Psychology*, 56, 805-814.

Bates, I. (1995) The Competence Movement and the National Vocational Framework: The Widening Parameters of Research. *British Journal of Education and Work*, 8, 2, 5 - 13.

Bates, I. (1999) A keynote paper presented at Researching Work and Learning, A First International Conference, School of Continuing Education, University of Leeds, UK.

Beaumont, G. (1995) *Review of 100 NVQs and SVQs. A Report submitted to the Department for Education and Employment.* DfEE: London.

Black, J. H., Hall, J., Martin, S. and Yates, J. (1989) *The Quality of Assessments: Case Studies in the National Certificate*. Scottish Council for Research in Education: Edinburgh.

DiBella, A. J., Nevis, E. C. and Gould, J. M. (1996) Understanding Organizational Learning Capability. *Journal of Management Studies*, 33, 3, 361-379.

DfEE (2000) Department for Education and Employment Research Programme 2000-2001 www.dfee.gov.uk/research/2000_2001.htm.

Employment Department (1994) *Thinking and Learning at Work, Core Skills and Training for Transfer*. Employment Department Group: Sheffield.

Eraut, M, and Steadman, S. (1998) *Evaluation of Level 5 Management S/NVQs Final Report 1998*. Research Report Number 7. University of Sussex at Brighton.

Eraut, M., Steadman, S., Trill, J. and Parkes, J. (1996) *The Assessment of NVQs*. Research Report Number 4. University of Sussex at Brighton.

Hammersley, M. (2000) *Diversity or Control in Educational Research*, A paper presented at the conference Diversity or Control in Educational Research, City University, Northampton Square, London, 27 January 2000.

Hodgson, A. and Spours, K. (2000) *Qualifying for Success: Towards a framework of understanding*. the first report from the project Broadening the Advanced Level Curriculum: Institutional Responses to Qualifying for Success, Institute of Education/ Nuffield Project: London.

Jackson, N. S. and Jordan, S. S. (1999) *Skills Training: Who benefits?* Conference proceedings of Researching Work and Learning. A First International Conference, School of Continuing Education, Trinity and All Saints College, Leeds, September 10-12.

Konrad, J. (1998) *Assessment and Verification of National Vocational Qualifications: a European quality perspective*. Education-line www.leeds.ac.uk/educol/index.html

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

QCA (1999a) *Unitisation and credit in the national qualification framework: A position paper*. QCA: London.

QCA (1999b*) Improving NVQs: a report of the User Forums held by QCA in May and June 1999.* www.qca.org.uk.

QCA (1999c) *NVQ Monitoring Report 1998/1999*. www.qca.org.uk/nvq-monitoring/98-summ.htm.

QCA (2000a) *Finding your way around. A leaflet about the National Qualification Framework.* QCA: London.

QCA (2000b) *Annual NVQ Statistics Supplement 2000*. DataNews. www.qca.org.uk

QCA (2000c) *QCA Website*. www.qca.org.uk.

Raggatt, P. and Williams S., (1999) *Government, Markets and Vocational Qualifications. An Anatomy of Policy*. Falmer Press: London.

Robinson, P. (1996) *Rhetoric and Reality: Britain's New Vocational Qualifications.* Centre for Economic Performance, London School of Economics: London.

Scharaschkin, A. and Baird, J. (2000), The Effects of Consistency of Performance on A Level Examiners' Judgements of Standards, *British Educational Research Journal*, 26, 3, 343 - 358.

Senker, P. (1996) The Development and Implementation of National Vocational Qualifications: an engineering case study. *New Technology, Work and Employment*, 11, 2, 83 - 95.

Smithers, M. (1994) *All Our Futures Britain's Education Revolution*. A Dispatches Report on Education. Channel Four Television: London.

Stephenson, J., Williams, R., Cairns, L. and Critten, P. (1999) *The Contribution of National Vocational Qualifications (NVQs) to the Learning Milieu of the work-place.* A research report for RSA Examinations Board: Coventry.

Williams, R., Cunningham, L., and Stephenson, J. (1997) *The use of NVQs as a means to develop Corporate Capability*. A research report for RSA Examinations Board: Coventry.

Wolf, A. (1995) *Competence Based Assessment*. Open University Press : Milton Keynes.

Wolf, A. (1998) Portfolio assessment as national policy: the National Council for Vocational Qualifications and its quest for a pedagogical revolution. *Assessment in Education, Policy and Practice*. 5, 3, 413-445.

Acknowledgement

Sonia Osborn and Stephen Hunt from OCR have kindly offered their time and knowledge which were valuable in writing this paper. Although my colleagues Sonia and Stephen have contributed greatly to this paper they do not necessarily agree with the views of the author.