

# **Research and Development Series**

The Potential for Computer Assisted Assessment in the  
Assessment of National Vocational Qualifications

December 1991

Report No.6

**A Technical Report published by the Employment  
Department Standards Methodology Branch.**

Guildford Educational Service Ltd

The Potential for Computer Assisted Assessment in the  
Assessment of National Vocational Qualifications

Final Report

Report to the Employment Department - TEED

Project Manager: Christine Ward

Guildford Educational Services Ltd  
32 Castle Street  
Guildford GU1 3UW

November 1991

Report to Employment Department - TEED

on the

Potential for Computer Assisted Assessment in the  
Assessment of National Vocational Qualifications

CONTENTS

VOLUME I: REPORT

	Page
<b>Part I: Introduction</b>	<b>1</b>
<b>SECTION 1: THE PROJECT</b>	<b>1</b>
Terms of reference	1
Background	1
Methodology	2
General National Vocational Qualifications	2
Outcomes	3
Staffing	3
Acknowledgements	3
<b>SECTION 2: METHODOLOGY AND SURVEY RESULTS</b>	<b>4</b>
Introduction	4
Awarding and industry bodies	4
Centres	4
TECs	4
Software	5
Research	5
Other bodies	5
Results of survey of centres	6
Results of industry body survey - general	6
Results of survey of TECs	7
Other perspectives	7
Software currently in use	8
The demands made by NVQ assessments	8
Coverage of NVQ assessments	9
Clarity of information and guidance	10
Assessors and verifiers	11
Recording and reporting assessment results	11
Equipment available in centres	11
Assessment resources	12
Priorities for action	12

	Page
<b>Part II: IT Application in the field of assessment</b>	13
<b>SECTION 3: OVERVIEW OF THE USE OF INFORMATION TECHNOLOGY IN THE ASSESSMENT PROCESS</b>	13
Introduction	13
Main features of each system	14
Contribution to access, efficiency and effectiveness	14
Current availability and use	15
On-going research	15
Potential for further development	15
Possible markets	15
Possible follow-up action	17
<b>SECTION 4: CENTRAL RECORDING AND PROCESSING OF ASSESSMENT DATA</b>	19
Main features	19
Contribution to improving efficiency and access	19
Current availability and use	20
On-going research	20
Potential for further development	20
Possible markets	21
Possible follow-up action	21
<b>SECTION 5: LOCAL RECORDING AND TRACKING OF ASSESSMENT</b>	22
Introduction	22
Recording achievement of elements - main features	22
Contribution to access, efficiency and effectiveness	22
Recording achievement of elements - Current availability and use	23
Recording achievement of elements - On-going research	23
Recording achievement of elements - Potential for further development	24
Recording achievement of elements - Possible markets	24
Recording achievement of elements - Possible follow-up action	25
Individual Action Plans - Main features	25
Contribution to access, efficiency and effectiveness	25
Individual Action Plans - Current availability and use	26
Individual Action Plans - On-going research	26
Individual Action Plans - Potential for further development	27
Individual Action Plans - Possible markets	28
Individual Action Plans - Possible follow-up action	28
Systems for use by candidates	28
<b>SECTION 6: TEST DESIGN AND PRODUCTION</b>	30
Main features	30
Contribution to access efficiency and effectiveness	31
Current availability and use	31
On-going research	32
Potential for further development	32
Possible markets	33
Follow-up action	34

	Page
<b>SECTION 7: COMPUTER DELIVERY OF ASSESSMENT - GENERAL</b>	<b>35</b>
Introduction	35
Contribution to improving access, efficiency and effectiveness	36
Current availability and use	38
On-going research	38
Potential for further development	38
Possible markets	39
Possible follow-up action	40
<b>SECTION 8: DELIVERY OF SKILLS ASSESSMENT</b>	<b>43</b>
Main features	43
Contribution to improving access, efficiency and effectiveness	43
Current availability and use	44
On-going research	45
Potential for further development	45
Possible markets	48
Possible follow-up action	48
<b>SECTION 9: DELIVERY OF KNOWLEDGE ASSESSMENTS</b>	<b>50</b>
Main features	50
Contribution to improving access, efficiency or effectiveness	50
Current availability and use	50
On-going research	51
Potential for further development	52
Adaptive testing	54
Possible markets	56
Possible follow-up action	56
<b>SECTION 10: COMPUTER SIMULATION</b>	<b>57</b>
Main features	57
Contribution to improving access, efficiency or effectiveness	57
Current availability and use	58
On-going research	59
Potential for further development	59
Possible markets	61
Possible follow-up action	62
<b>SECTION 11: AIDS TO DECISION-MAKING</b>	<b>63</b>
Main features	63
Contribution to improving access, efficiency or effectiveness	63
Current availability and use	64
On-going research	64
Potential for further development	65
Possible markets	66
Possible follow-up action	66

	Page
<b>SECTION 12: ACCREDITATION OF PRIOR LEARNING</b>	67
Main features	67
Contribution to improving access, efficiency or effectiveness	67
Current availability and use	67
On-going research	68
Potential for further development	68
Possible markets	69
Possible follow-up action	69
<b>SECTION 13: ASSESSORS AND VERIFIERS</b>	70
Introduction	70
Main features	70
Contribution to improving efficiency and access	70
Current availability and use and on-going research	70
Potential for further development	71
Possible markets	71
Possible follow-up action	71
<b>SECTION 14: FEEDBACK AND QUALITY</b>	72
Main features	72
Provision for improving access, efficiency or effectiveness	73
Current availability and use	73
On-going research	74
Potential for further development	75
Possible markets	76
Follow-up action	76
<b>Part III: Conclusions and Recommendations</b>	77
<b>SECTION 15: CONCLUSIONS OF PROJECT AND CRITERIA FOR ADOPTING CAA</b>	77
Introduction	77
Types and characteristics of CAA applications	77
Contribution of CAA to improving access, efficiency and effectiveness	78
Potential disadvantages of CAA	80
Criteria for adopting CAA	80

VOLUME II: ANNEXES

	Page
Annex A Full project specification	AX 1
Annex B List of visits	AX 5
Annex C Publications consulted	AX 7
Annex D Results of surveys	AX 13

Guildford Educational Services Ltd

Report to Employment Department - TEED

on the

Potential for Computer Assisted Assessment in the  
Assessment of National Vocational Qualifications

Part I: Introduction

SECTION 1: THE PROJECT

Terms of reference

- 1.1 Guildford Educational Services Ltd (GES) was contracted by the Training, Enterprise and Education Division of the Employment Department (TEED) to conduct a survey of the potential for computer assisted assessment of National Vocational Qualifications (NVQs). The project commenced on 1 April 1991 and reported in mid-October 1991.
- 1.2 The aim of the project was to investigate the potential for computer assisted assessment (CAA) for improving access, efficiency and effectiveness of assessment for accreditation to nationally recognised standards of competence and to inform the Department of possible interventions in the development and implementation of CAA. The full terms of reference are given in Annex A.

Background

- 1.3 The overall background to the project had two main aspects, one of which was the developing framework of National Vocational Qualifications based on industrially defined standards of competence. Standards of competence are being developed by Lead Bodies as part of Employment Department's Standards programme, and qualifications based on those standards are accredited as NVQs by the National Council for Vocational Qualifications. This process is leading to the development of a more comprehensive and flexible framework of qualifications than was available in the past, with emphasis on the standards attained rather than the educational or training process leading to the qualification.
- 1.4 Within the NVQ system, emphasis is placed mainly on the practical demonstration of competence in the workplace, but there is a growing realisation that workplace observation needs to be supplemented by other assessment methods. Reasons put forward for this include:
- concerns about the quality of assessment, in terms of both coverage and consistency
  - the difficulty of covering the whole range of situations in which candidates are expected to be competent by means of workplace observation



- the need to assess underpinning knowledge and understanding which cannot necessarily be inferred from workplace performance
  - pressures for increased rigour in the assessment process
  - the heavy burden placed on assessors, both in assessment and in recording
  - the cost of assessment
  - difficulties of access for candidates not in employment or employed in an occupation with a restricted range of opportunities to demonstrate competence.
- 1.5 Other problems were believed to be occurring because of difficulty in interpreting the requirements of the standards for individual competence elements or for combinations of evidence and because of difficulties in coordinating the standards of assessors.
- 1.6 One part of the project was to conduct a survey of centres and Lead Bodies to discover how far their experience confirmed that such difficulties were being encountered in practice.
- 1.7 The second major aspect to the overall background to the project was the developments made in information technology during the previous decade. Such developments have been harnessed to provide imaginative computer based training packages, as well as interactive video and other more advanced applications, and it was thought that similar techniques might be used to solve some of the problems of assessment.

#### **Methodology**

- 1.8 As described in more detail in Section 2, the methodology for the project was a combination of software review, interviews and visits to both software producers and awarding and Lead Bodies, and surveys by means of questionnaires.

#### **General National Vocational Qualifications**

- 1.9 NCVQ's Consultation Paper on General National Vocational Qualifications (GNVQs) was published in early October 1991, too late to have any impact on the conduct of the project.
- 1.10 If GNVQs are developed along the lines set out in the Consultation Paper, with externally set assessments of underpinning knowledge on a fairly restricted number of units, they will increase the potential market for computer-assisted assessment. However, larger numbers of candidates for assessment in GNVQ units would necessitate the availability of additional hardware in the centres if CAA is to be used.
- 1.11 It is recommended that, when the proposals for GNVQs are finalised, Employment Department and/or NCVQ should commission a brief supplementary discussion paper on the role which computer assisted assessment could play in GNVQs.

## Outcomes

- 1.12 The major outcome of the project is this report, which surveys existing and potential developments in computer assisted assessment, relating them to identified needs and to the contributions which they can make to the improvement of access, efficiency and effectiveness.
- 1.13 The report also identifies further development which is needed to make computer assisted assessment techniques fully operational. Outline specifications have been prepared for projects which it is suggested that the Employment Department might wish to undertake in order to promote the development and implementation of computer assisted assessment.

## Staffing

- 1.14 The Project Manager was Christine Ward, working under the general direction of John Twining, Chairman of Guildford Educational Services. Administrative support, particularly in relation to the surveys, was undertaken by Robin Twining, and secretarial support provided by Blue Stevens.

## Acknowledgements

- 1.15 The project team would like to express their thanks to software producers, awarding and Lead Body staff, TEC staff and staff at centres who responded to their requests for information and completion of questionnaires, and took time to discuss with the project team issues relating to computer assisted assessment.
- 1.16 The team would also like to express their thanks to members of the Steering Group for their very positive contributions to the conduct of the project and to Dr William Ward of ETS who supplied information, advice and encouragement from the US viewpoint and made an important contribution to the production of the report and associated project specifications.
- 1.17 Thanks are also due to David Stuart and David Porter of the Employment Department and David Bell (External Project Manager), for their support and assistance.

## SECTION 2: METHODOLOGY AND SURVEY RESULTS

### Introduction

- 2.1 This section provides an outline of the methodology used in the project, together with a summary of the results obtained from the survey questionnaires and a discussion of the issues arising. Supporting information is given in the Annexes.

### Awarding and industry bodies

- 2.2 Contact was made with all the major vocational awarding bodies in England and Wales to identify what activities already existed in relation to computer assisted assessment and what interest they might have in CAA. Discussions were held with two City and Guilds staff members and, more briefly, with staff at LCCI and Pitman Examinations Institute. A list of all visits undertaken, both to awarding bodies and to other organisations, is given in Annex B.
- 2.3 Questionnaires were sent to the Industrial Training Organisations and Lead Bodies on the Employment Department's list, asking them to indicate their current involvement, if any, in computer assisted assessment, areas which they were interested in developing, and problems which they were encountering or envisaged they might encounter in the assessment of the NVQs with which they were concerned. Completed questionnaires were received from 55 industrial bodies.
- 2.4 Follow-up discussion by telephone was undertaken with some of those who responded to the questionnaire.

### Centres

- 2.5 Questionnaires were despatched to all LEA colleges in the UK, to all YT and ET managing agents on the latest available list, and to the personnel departments of the 200 largest firms in the UK. Responses were received from 250 organisations, of whom 217 were colleges. The survey asked centres to indicate what problems they were encountering in the assessment of NVQs, what equipment they had available for possible use in candidate assessment, and what resources they would like to see made available for use in assessment.
- 2.6 A limited amount of telephone follow-up to the questionnaire was possible, and two centres were visited.

### TECs

- 2.7 Preliminary letters were sent to all TECs and LECs asking them if they had an interest in the potential for computer assisted assessment, and if they were willing to complete the questionnaire. Most (32) of the 45 who responded indicated that they did not feel able to complete a questionnaire at the present time, but expressed an interest in being kept informed of developments. Although 9 TECs asked for a questionnaire, only 5 completed questionnaires were returned.

- 2.8 GES has played a small part in the development of the training credits pilot by SOLOTEC (the South London Training and Enterprise Council), which uses a smart card system. As well as providing an illustration of one use of technology in recording the attainment of NVQs and units, contact with participants in the training credit pilot provided interesting insights into the difficulties which centres were encountering in conducting NVQ assessments.

#### **Software**

- 2.9 An initial letter asking for information about CBT and other technology-based training products, was sent to open learning producers identified as being possibly involved in this area. The mailing list was derived mainly from the Open Learning Directory, with the addition of other organisations known to the project team. As far as possible, demonstration discs of interesting products were obtained and viewed in-house. Brief information is provided about software products viewed (Annex E) and about on-going developments in Annex F.
- 2.10 A small number of software producers were visited. Visits to exhibitions (specifically the Human Resources Development Exhibition and the Education, Training and Personnel Development Exhibition), and a visit to the Employment Department's LTU Learning Resource Centre, provided opportunities to view a wide range of CBT and related products. The exhibitions provided a more cost-effective way of viewing a selection of software than visits to individual producers, as did visits (undertaken in the course of other work) to the International Conference on Technology & Education, Toronto and the International Labour Office Conference, Tashkent.

#### **Research**

- 2.11 Background research included reading or re-reading a range of publications relating to NVQs and the Standards programme, and to developments in automated testing. A bibliography of the project is given in Annex C.
- 2.12 Visits to two members of the Steering Group, David Bartram at the University of Hull, and Charles Johnson at Psychometric Research and Development, provided an opportunity to discuss some of the issues in more detail. A visit to the Institute of Manpower Studies at Brighton University led to a stimulating (and challenging) discussion of the strengths and weaknesses of the NVQ system as seen from the outside.

#### **Other bodies**

- 2.13 Letters were sent to a small number of national bodies asking them for any information they had on current or proposed activities in computer assisted assessment. The replies overlapped with initiatives discovered from other information sources, but also provided additional useful information. The bodies contacted were the Further Education Unit, Further Education Staff College, National Council for Educational Technology and the Scottish Council for Educational Technology.
- 2.14 Dr William Ward of ETS supplied information on developments in the United States and provided advice throughout the project, but especially in relation to the preparation of the report and the proposed follow-up projects.

## Results of survey of centres - general

- 2.15 Survey questionnaires were sent to 1209 centres. A total of 250 forms were returned:
- 217 from educational establishments
  - 29 from YT/ET managing agents
  - 11 from companies
  - 3 unclassified.
- 2.16 A number of educational establishments and some companies were also YT/ET managing agents; in such cases they were classified as education or company. Some educational establishments also submitted multiple returns.
- 2.17 Annex D provides a copy of the questionnaire, together with a frequency count of the number responding to each question, including information about how many colleges responded to each question. Not all respondents completed all questions.
- 2.18 Annex D also contains details of the occupational areas mentioned in the completed questionnaires. Many respondents indicated that they dealt with more than one occupational area. A very wide spread of occupations was covered overall, but by far the most common area quoted was Business Administration. Most responses were made on the basis of experience of running existing NVQs: some centres reported that NVQs were still being introduced in their area of interest and they could therefore only anticipate what the problems would be.
- 2.19 Centres responded enthusiastically to the opportunity to itemise the difficulties they were encountering in NVQ assessment. Overall, companies and YT/ET managers tended to find fewer difficulties or 'scope for improvement' than the educational establishments. In general, also, the tone of comments indicate a developing interest in the use of IT in assessment.
- 2.20 Subsequent paragraphs (2.29 onwards) discuss centre responses in more detail, and compare them with the responses of the industry bodies and with comments received from elsewhere. The major issues to emerge from the centre survey were those of assessor time, cost of assessment and the amount of paperwork involved.

## Results of industry body survey - general

- 2.21 Of the 55 industrial organisations which completed the questionnaire, 50 were Lead Bodies; 40 of the bodies which responded also had a role as awarding bodies, 9 of them acting individually and 31 as partners in joint awarding bodies. About half the industrial organisations which completed the questionnaire were concerned with standards which were already being implemented for at least one NVQ level. The other bodies mainly had standards in preparation or standards which were completed but had not yet been implemented.

- 2.22 The questions asked of the Lead Bodies were not identical to those asked of centres. The responses indicated that for industry bodies the major issues were those of assessor time and cost of assessment, the need to cover range statements and rarely occurring situations, and the need to monitor centres.
- 2.23 In addition to the information obtained from completed questionnaires, useful discussions were held with the Royal Institute of Public Health and Hygiene and the Associated Examining Board, and contact was made with the British Polymer Training Association.

#### Results of survey of TECs

- 2.24 Because so few TECs completed questionnaires, no meaningful analysis is possible. Where the TECs had relevant experience of operating NVQs their responses reinforced those received from centres and industry bodies. Discussions with SOLOTEC provided more detailed information about problems being encountered in the Training Credits Pilot.
- 2.25 Qualitec (St Helens) reported that it was setting up an assessment centre for clients entering TEC funded programmes, and also had an interest in offering NVQ assessment. Several TECs which made contact with the project indicated that they were interested in taking part in the Employment Department's programme relating to the Accreditation of Prior Learning.

#### Other perspectives

- 2.26 Discussions with awarding bodies, experienced standards consultants, the Institute of Manpower Studies and NCVQ and Employment Department staff acknowledged the problems of cost and assessor time but laid much greater emphasis on the need for **quality** of assessment. This has two major facets, coverage and consistency.
- 2.27 Adequate coverage of assessment was seen as being difficult to achieve by workplace observation alone, since the workplace might not provide an adequate range of experiences or the opportunity to demonstrate competence in the full range of circumstances. Workplace observation would not necessarily provide sufficient evidence of the candidate's underpinning knowledge and understanding.
- 2.28 The diversity of work settings and the number of workplace assessors involved could also lead to problems in achieving **consistency** of assessment. There was a strong need for intervention to improve the consistency of the judgements made by assessors, and it was suggested that newer awarding bodies might not appreciate the importance of measures such as assessor training and guidance material for assessors. The overall consistency of the assessment process could also be improved if workplace observation were supplemented by more easily standardised methods.

## Software currently in use

- 2.29 Annex D provides a summary of the responses given to Question 4 of the centre survey, which asked for information about software currently being used for assessment purposes. A number of centres were using software to monitor candidate progress, Rocket being the most commonly quoted example. Many centres were also making use of standard proprietary wordprocessing, spreadsheet and database software to assess candidate competence in the use of this software. Proprietary packages for keyboard training were also used for assessment purposes. There was evidence of a small amount of use of subject specific knowledge tests and also of software packages which allow computer delivery of knowledge tests in any subject area.
- 2.30 A small number of industry bodies reported that they already had some form of computer assisted assessment. The most commonly quoted application was computerised record keeping at national level, but seven bodies also reported that they had computerised record keeping at centre level. There were also four references to computer delivered tests of underpinning knowledge, two of computer delivered skills tests, six of computer simulations of fault-finding, and three of plant operation. It was not possible within the timescale of the project to follow-up all these references.

## The demands made by NVQ assessments

- 2.31 Centres indicated that the major problem in the demands made by NVQ assessment was the amount of assessor time required; 53% of centres reported 'considerable problems' in this area, with a further 26% reporting 'minor problems'. There was also considerable concern about the cost of assessment, with 41% of centres reporting 'considerable problems' and 24% reporting 'minor problems'. Lower numbers reported concern with candidate time, and the use of equipment, and written comments referred to the complexity of the NVQ system and the amount of paperwork required.
- 2.32 Amongst industry bodies the proportion reporting 'considerable problems' with assessor time and cost was similar to the results from centres, but when numbers reporting either 'considerable' or 'minor' problems in these areas are considered together, the figures rise most noticeably; 85% of industry bodies reported considerable or minor problems relating to the amount of assessor time required, and 82% reported problems in relation to the cost of assessment.
- 2.33 The concern with the amount of assessor time is confirmed by other written and oral comments received in the course of this and other projects. The major difficulty seems to be encountered by colleges and managing agents who need to arrange assessment for candidates during work placement periods. Small firms (who may provide the majority of work placements) are reluctant to become involved in the assessment process where they only have one trainee to deal with and have no trained assessors. Staff from the college or managing agent therefore need to go out to undertake assessment in the workplace for each individual trainee, and it is estimated that this may sometimes require 15 hours for each trainee. Time is also needed to gain employer cooperation for the process.

- 2.34 Estimates produced by a managing agent in the SOLOTEC area which have been made available to GES suggest that the cost of assessment for a candidate on work placement can be in the order of £200-£250 per trainee per year for Business Administration, and around £500 per trainee per year in Caring. For Construction the costs are estimated at £50-£100 per trainee per unit. These costs include fees to awarding bodies and costs of materials, but assessor time can be a major factor, especially where the assessor comes from outside the organisation.
- 2.35 SOLOTEC staff also reported that the difficulty of finding suitable work placements for trainees during a recession was being exacerbated by employers' fears that they would be required to spend substantial time in assessing the trainees. When work placements have been lost it is difficult to recover them or to find alternatives.
- 2.36 In response to the question about problems in NVQ assessment, some centres drew attention to concerns over reliability and consistency, both between the centres themselves and between awarding bodies basing awards on the same standards. This led to some concern about the credibility of NVQs, which was reinforced by discussions with other organisations.

#### Coverage of NVQ assessments

- 2.37 Compared with cost and assessor time, fewer centres reported 'considerable difficulty' in arranging assessments to cover all the elements and all the range statements and to deal with rare situations, dangerous and emergency situations and the need for assessment in the workplace; between 17% and 27% of centres reported 'considerable difficulty' in relation to these factors, with 32%-48% reporting 'minor difficulty'.
- 2.38 Written comments indicated that the workplace was sometimes unable to provide facilities for all the aspects of assessment required by the standards. For example, some companies lacked the full range of necessary equipment or the full range of types of work.
- 2.39 Considerations of commercial confidentiality sometimes made companies reluctant to allow an outside assessor or verifier full access to candidate's work.
- 2.40 The coverage of assessments for NVQs was a more serious concern for industry bodies, with around 30% recording considerable problems under this heading, and up to 82% reporting either considerable or minor problems. The area which provoked most concern was arranging assessment of situations which rarely occur in practice.
- 2.41 Industry Bodies were also asked whether they were encountering problems in arranging assessment for candidates who had not followed a conventional course or training programme. Only 18% regarded this as a considerable problem, but an additional 33% regarded it as a minor problem.



## Clarity of information and guidance

- 2.42 Centres were asked to indicate the level of problems they were experiencing in interpreting the information provided on elements, performance criteria and assessment requirements, and in deciding candidates results. No comparable questions were asked of Industry Bodies.
- 2.43 The percentage of centres considering that there was 'considerable scope for improvement' in the wording and clarity of elements and performance criteria was 26, with 42% indicating that some improvements were needed; 22% thought that there was considerable scope for improvement in the definition of the range of tasks or situations required, and 33% that there was considerable scope for improvement in the guidance as to the difficulty or complexity of the tasks expected.
- 2.44 In response to these questions, several centres reported that the wording of the standards was difficult to understand, particularly for candidates, and that the documentation was not user-friendly.
- 2.45 Some written comments indicated a difficulty in relating the new standards to known benchmarks; it was suggested that an indication of the relationship of standards to RSA Typewriting 2 and to a statement of typing speed in words per minute would be useful. There was also a plea for provision of sample assignments and for more written tests.
- 2.46 It is apparent from some of the comments received informally that there are severe problems of dissemination of information similar to those which have accompanied other major developments in vocational education and training in the past. For example, several centres were anxiously but erroneously anticipating that new standards for Construction would be published for implementation in September 1991. There is also considerable confusion over the assessment arrangements and the status of assessing and awarding bodies in the Care sector.
- 2.47 In response to questions about the difficulty of determining candidate results, relatively small percentages of centres reported 'considerable difficulty'. However, exactly half the centres reported either considerable or minor difficulty in deciding whether candidates had fulfilled the requirements for a particular unit or element, and 55% reported either considerable or minor difficulty in weighing the value of different types and combinations of evidence.
- 2.48 Written comments suggested that some centres were feeling the lack of trained and supportive verifiers, and others considered that it would be useful to be able to compare notes with other centres.

### Assessors and verifiers

- 2.49 Centres did not report serious difficulties in finding enough suitably qualified assessors or in coordinating the standards of assessors within the organisation, although it is notable that colleges had greater problems coordinating the standard of internal assessors than did other types of centre. However, 36% of centres reported considerable difficulty in coordinating the evidence of outside assessors such as workplace supervisors, and a further 28% reported minor difficulty. Written comments indicated that it was also very difficult to appoint assessors outside the organisation and that workplace supervisors were either inadequately trained or unwilling to participate in the assessment process.
- 2.50 Industry bodies reported slightly greater problems in finding enough suitably qualified assessors, with 35% indicating that they had considerable problems in this area, and a further 32% reporting minor problems.
- 2.51 The centres commented that the requirement for assessor training varied considerably between awarding bodies, vocational areas and subjects, and that where assessor training was required the cost involved was high.
- 2.52 Almost one-third of industry bodies reported considerable problems in finding enough suitably qualified verifiers, and this is reflected in their concerns over coordinating the standards of assessors, with 78% recording either considerable or minor problems in this respect. Centres did not see liaison with awarding body verifiers as an area of difficulty, although a few commented on difficulties with individual awarding bodies and the disadvantage of being charged extra for additional verifier visits. One centre described arrangements for liaising with the City and Guilds verifier as 'excellent'.

### Recording and reporting assessment results

- 2.53 The administrative procedures involved in recording and reporting candidate results were a serious problem area for many centres, with 38% reporting 'considerable scope for improvement' in systems for recording candidate performance, and 46% reporting 'considerable scope for improvement' in handling the associated paperwork. A number of centres added written comments to reinforce this point, one stating that the volume of paperwork was 'prodigious'.
- 2.54 Several of the centres adding written comments made a plea for a common national proforma for recording candidate results for all subjects and awarding bodies. Several also commented on the lack of support staff to help assessors with the paperwork.

### Equipment available in centres

- 2.55 The responses of centres to questions about the available equipment indicated that computer assisted assessment could proceed on the assumption that at least three-quarters of centres have an IBM compatible microcomputer with hard disc available for use in the assessment process. However, the number of machines available in each centre is likely to be limited and this will have implications for the way in which computer delivered assessment can be organised.

- 2.56 The proportions of centres with more advanced equipment were lower, with only 11% having a CD-ROM player, 15% having interactive video equipment, and 2% having DVI equipment.

#### **Assessment resources**

- 2.57 In response to a question asking what assessment resources they would like to see made available for use in NVQs, centres particularly highlighted the desirability of computerised systems for recording and monitoring candidate progress and computerised tests of underpinning knowledge; each of these was supported by 82% of centres. Other suggestions which proved popular were computerised tests of knowledge or understanding designed to assess students or trainees entering the course and banks of paper based assignment or assessment material. The popularity of computer delivered skills tests (45%) and computer simulations (44%) was lower, probably reflecting the fact that these are not relevant across all subject areas.
- 2.58 By contrast, the industry bodies did not attach a high priority to the development of CAA applications. The aspects which Industry Bodies indicated that they would most like to develop were computerised record keeping at national level (36% of Bodies) and at centre level (31%), followed by computerised guidance for APL (29%) and computer simulations of reaction to dangerous or emergency situations (27%). Only 20% of industry bodies gave a high priority to developing computer delivered tests of underpinning knowledge.

#### **Priorities for action**

- 2.59 The results of the surveys show that the primary concern of the centres is with the immediate problems of cost and resources, including time. These need to be addressed, whether by computer assisted assessment or by other means. However, attention also needs to be devoted to the more fundamental and long term issues of assessment quality (particularly consistency and breadth of coverage) which were emphasised more strongly by industry and awarding bodies and by Employment Department staff.

## Part II: IT Application in the field of assessment

### SECTION 3: OVERVIEW OF THE USE OF INFORMATION TECHNOLOGY IN THE ASSESSMENT PROCESS

#### Introduction

- 3.1 Sections 4-14 of this report describe existing and potential applications of information technology in the assessment process. Each section is organised under the main headings of:
- the main features of the application
  - contribution to improving access, efficiency and effectiveness
  - current availability and use
  - on-going research
  - potential for further development
  - possible markets
  - possible follow-up action.
- 3.2 This section provides an introduction to Sections 4-14, and highlights some general issues and topics, together with some general points related to each heading.
- 3.3 Each section of the report focuses on an application of information technology within the assessment process, and may therefore relate to more than one type of technology. For example, interactive video is mentioned under assessment of knowledge and also under simulations.
- 3.4 Since the major focus of the project was on assessment for National Vocational Qualifications related to industrial standards, the report deals mainly with summative assessment. Formative assessment, although an important topic, and one which was shown by the survey to be of considerable interest to centres, was felt to be of less central importance. However, many of the techniques described could be used for formative as well as for summative assessment.
- 3.5 In the NVQ system where the emphasis is on evidence of competence, the dividing line between formative and summative assessment may sometimes be blurred. There appears to be no reason why an assessment intended mainly for formative purposes should not be used as evidence of competence in the summative assessment if the candidate has fulfilled all the requirements of one or more elements. Conversely, a candidate who fails an end-of-course assessment should be entitled to feedback on areas of weakness as a contribution to his or her further training. Accreditation of Prior Learning can have both a formative and summative role.

### Main features of each system

- 3.6 Under this heading, each of Sections 4-14 describes the main features of the information technology application under consideration, including systems which are currently in use for assessment at the present, and potential uses of IT in assessment.

### Contribution to improving access, efficiency and effectiveness

- 3.7 Under this heading, each of Sections 4-14 discusses the contribution which the IT application under consideration could make to the assessment process in terms of the main strands of access, efficiency and effectiveness.
- 3.8 This sub-section also refers to problems in assessment identified by the market survey, highlighting the contribution which each application could make to solving those problems.
- 3.9 **Access** refers to the ease with which candidates can obtain qualifications or can be assessed as a contribution towards a qualification. One of the major principles of the NVQ system is that there should be no unnecessary barriers to access. Issues of access relate particularly to non-standard candidates, for example minority groups, the disabled, those who have not attended a formal course, those who are not in relevant employment and (following a recent announcement by the Secretary of State for Education and Science) candidates in prisons. However, the emphasis on workplace assessment may also create barriers for candidates who are following a college course in conjunction with employment which does not cover the whole range of the Standards. Access will be improved if alternatives to workplace assessment are more readily accepted and if any formal assessments used are available at frequent intervals.
- 3.10 **Efficiency** relates not only to the cost of assessment in direct monetary terms, but also to the resources, including assessor time, which need to be devoted to the assessment process. As described in Section 2, cost and assessor time were two of the three major concerns of centres in relation to NVQs.
- 3.11 The **effectiveness** of the assessment process is mainly a matter of its quality. One aspect of quality is the appropriateness of the content and coverage of the assessment, and this report considers, for example, how computer delivered assessment could use simulations to assess situations which occur only rarely or are difficult to assess in other ways. This report also looks at the contribution of CAA to extending the breadth and range of the assessment process and its possible contribution to the assessment of underpinning knowledge and understanding. Issues of quality and effectiveness also relate to the consistency with which the standards are applied, for example consistency between different centres or assessors, and the need to avoid 'contamination' where the result is affected by factors which are irrelevant to the competence of the candidate.

- 3.12 Issues of access, efficiency and effectiveness are often inter-related. For example, if computer assisted assessment can improve efficiency and cut costs then this will effectively improve access by ensuring that more candidates or their employers can afford the assessment process.
- 3.13 The contribution of computer assisted assessment may be indirect. For example, CAA could make it easier for awarding bodies to introduce knowledge assessments without increasing overall costs or restricting access. The use of knowledge assessment may improve the effectiveness of the assessment by increasing its range.

#### **Current availability and use**

- 3.14 Under this sub-heading, Sections 4-14 describe actual uses of technology within assessment, for NVQs and other qualifications. This sub-heading also describes techniques used in other areas, for example in computer based training, where these have the potential for transfer to assessment. In many sections the actual uses of the techniques for assessment purposes are rare.
- 3.15 Brief details of the software products mentioned under this heading are provided (Annex E).

#### **On-going research**

- 3.16 Under this sub-heading, Sections 4-14 refer to current research and also to current development projects in computer assisted assessment or related areas. More information about current developments is provided in Annex F.

#### **Potential for further development**

- 3.17 Under this sub-heading Sections 4-14 describe ways in which CAA applications could be made more effective or could be developed into other areas.
- 3.18 The discussion also covers ways in which technology not currently used for assessment could be adapted for assessment use, and in many sections of the report this is a more significant area for development.

#### **Possible markets**

- 3.19 There are several potential markets for different CAA products:
- candidates
  - centres
  - awarding bodies
  - other bodies
- 3.20 The candidate market is notionally a very large one (probably 2.5-3 million a year). Candidates do not themselves form a direct market, but their numbers are relevant to the markets of awarding bodies, centres and others who may provide an assessment resource. Table 1 gives the annual numbers of UK candidates being processed by the 6 major awarding bodies and two industrial awarding bodies (Engineering Training Association and Glass Training Ltd). The source is the GES report to the Employment Department on the National Qualifications Survey.

3.21 However, the candidate market is very fragmented. Table 2 shows the number of qualifications and units for some awarding bodies estimated from the NCVQ database and other information. Entries are likely to vary from under 20 to over 5000 per year for each unit, although many (probably most) candidates will take more than one unit in a year.

**Table 1: Number of candidates or registrations for processing**

	electronic form	paper form
BTEC	250K per year	-
C&G	550K per year	-
RSA	1M per year	2K per year
PEI	180K per year	9K per year
LCCI	140K per year	2K per year
SCOTVEC	250K per year	-
ENTRA	22K per year	-
GTL	-	500 per year

**Table 2**

	No of Quals	No of units/quasi-units
BTEC	200-300	100,000
C&G	600-700	1600-2000 *
LCCI	140-200	150-220 *
PEI	80-100	140-200 *
RSA	140-160	280-350 *
<b>Totals</b>	<b>1160-1460</b>	<b>2070-2770 + BTEC</b>

Total units on NCVQ database: 5400. Most of BTEC's claimed 100,000 units are centre-devised. Only a small proportion of BTEC-devised units are included in the NCVQ database.

\* Does not include units 'owned' by an ILB but assessed by C&G etc.

- 3.22 There are two main types of **centres** at which assessment takes place:
- a) those whose main mission is training and/or education (there are probably 1500-2000 of these, including colleges, schools, managing agents and private sector training organisations)
  - b) those which provide training and assessment for their own employees (there could be very large numbers of these, but many may only have one or two candidates a year, or perhaps less frequently).
- 3.23 Some 300 **awarding bodies** (including ITOs/LBs and professional bodies) are listed in the Qualification Aims volume of *Using Learning Information*. At present the number is growing as LBs realise ambitions to be awarding bodies. Whether many can remain viable on their own, and still provide a professional service to candidates; is questionable. Some rationalisation may therefore take place over the next few years.
- 3.24 **Other bodies** with an interest in providing an assessment service include TECs (82) and LECs (22). Their role may be to provide APL facilities or to facilitate workplace assessment by providing peripatetic assessors for centres with small candidate numbers.
- 3.25 In Sections 4-14 of the report, detailed discussion of possible markets for individual CAA developments is related to the results of the surveys and also indicates any subject areas which would be particularly suitable for specific CAA developments.
- 3.26 Some potential CAA applications are specific to one group of users, for example centres or awarding bodies. However, for computer delivered assessments (i.e. assessments which the candidate takes at the computer), marketing could take place in one of two distinct ways. Computer delivered assessment could be used by awarding bodies as a mechanism for administering the assessments which they set externally, in which case they would wish to be assured that the security and other procedures would be comparable to those in traditional uses of assessment. Alternatively, computer delivered assessment could be developed as an assessment resource for purchase by centres. Successful completion of a computer delivered assessment would then be an item of evidence of the candidate's competence which could be combined with other sources of evidence according to the needs of the individual candidate.

#### **Possible follow-up action**

- 3.27 Under this sub-heading, Sections 4-14 recommend action which the Employment Department could take or promote to develop the potential for CAA within NVQs. Outline specifications for the recommended follow-up projects are provided separately (Annex G).



- 3.28 Employment Department intervention may take one of a variety of forms, including:
- more research into what is available
  - a register of what is available
  - help in publicising existing products and on-going development work
  - work to set standards (e.g. for software) to facilitate the use of existing systems - for example in transferring data between centres and awarding bodies
  - work to define the infrastructure and security requirements for computer delivered assessment
  - development of 'shell' systems, which can be adapted for a range of subject areas
  - demonstrator projects producing subject-specific assessment material, as an example of what can be achieved technically and as an example of good practice
  - support for projects to identify the contribution which existing CBT products can make to assessment
  - other developments which might be subject-specific but would aim to explore issues of wider applicability.
- 3.29 Many of the proposals for follow up action draw on well established technology and/or assessment methodology and could be expected to produce useable results quite quickly. Others would explore newer technologies or assessment applications and would have less predictable outcomes.
- 3.30 Some suggested follow-up activities relate to more than one section of the report and some will be dependent upon earlier work. For example, Section 6 proposes the development of an assessment resource for Business Administration, which would incorporate computer delivered skills assessment as proposed in Section 8.
- 3.31 Sections 4-14 and the outline project specifications include recommendations for disseminating information about CAA in general and about specific CAA applications. It would also be helpful to produce a directory of CAA applications and the lists of packages and on-going research (Annexes E and F), although not suitable for publication in their present form, could provide the basis for this. Such a directory would be a useful information resource for anyone looking for 'off the shelf' products and for awarding and Lead Bodies considering adopting computer assisted assessment. An outline project specification for the production of such a directory is provided (G3.1).
- 3.33 The Employment Department might also wish to consider setting up a central support unit for CAA users, capable of providing software support and also responsible for the marketing of CAA products developed by organisations which do not have their own marketing resources.

## SECTION 4: CENTRAL RECORDING AND PROCESSING OF ASSESSMENT DATA

### Main features

- 4.1 Computer systems for central recording and processing of assessment data can play a major role in the tasks which any awarding body has to undertake:
- collecting details of students on registration or candidates on entry
  - in traditional examination systems, ensuring that the right number of the right question papers are sent to the right centre
  - sometimes keeping records of student/trainee progress
  - processing results of assessments and making decisions as to who should receive awards
  - printing and despatch of awards
  - maintaining records
  - providing or publishing statistics as required.
- 4.2 In a number of cases provision has been made for two or more awarding bodies to make awards jointly. In the NCVQ system this seems likely to increase.
- 4.3 Within the NCVQ system there is the additional possibility of the same candidate making up a complete NVQ with units achieved under the auspices of different awarding bodies.

### Contribution to improving access, efficiency and effectiveness

- 4.4 Any large-scale operation for handling awards requires computerisation. Although one awarding body with 500 candidates a year is known to use manual systems only, even at that level a computerised system would improve efficiency.
- 4.5 It is not necessary to computerise every single feature listed in paragraph 4.1. Some aspects, particularly the maintenance of records and the publication of statistics, are better supported by computerisation, even when numbers are comparatively low. Details of students/candidates and the processing of assessments probably become economic with as few as 100 candidates.
- 4.6 The main aim of computerising centralised recording and processing of assessment data is to improve efficiency by reducing manpower and costs. Computerisation should also improve the service provided to candidates and centres, for example by allowing a much quicker turn-round of results and certificate production. Computerised central recording and processing systems have no effect on the effectiveness of the assessment itself and only a marginal impact on access.

#### Current availability and use

- 4.7 All major awarding bodies have their own computer systems, usually on mainframes. However, not all these systems do the same thing or, if they do, do it in the same way. There is also a lack of standardisation across systems (for example in the number of characters allocated to candidate name and whether surname and forenames are treated as a single field). The differences between awarding bodies can cause difficulties for centres dealing with more than one body. Without additional work it could be difficult for awarding bodies to share data.
- 4.8 NCVQ has commissioned a company called Lodestar to develop a computer system which runs on IBM PCs for use by smaller awarding bodies. This is a 'shell' system but its customisation for individual bodies requires some degree of programming expertise.
- 4.9 A proprietary system called Hamlet has been sold to several GCSE and membership bodies, and would probably be suitable for adaptation for NVQ bodies, but again would need customisation for each individual use. Other systems may exist, but have not been identified.

#### On-going research

- 4.10 GES recently undertook a feasibility study on a National Qualifications Survey (NQS) for the Statistical Services Division of the Employment Department. This is now about to be followed up by a pilot survey which, if successful, could lead to standard candidate data being collected by all awarding bodies. As part of the National Qualifications Survey NCVQ has been asked to coordinate work with the major awarding bodies to see if there could be national adoption of a single candidate number for vocational qualifications. This would make it easier for a candidate to obtain parts of a single award from different awarding bodies as well as enabling the automatic tracing of a candidate's progression and the aggregation of the progression of different candidates to identify patterns.

#### Potential for further development

- 4.11 Few, if any, of the computer systems of the main awarding bodies provide the full flexibility of data handling required for NVQs/SVQs and their units. Most awarding bodies have plans for on-going internal development. To some extent this may be shared in the near future by the National Qualifications Survey. There is a lack of compatibility between existing systems, but the development of NQS may go some way to curing this.
- 4.12 There is considerable scope for developing the NCVQ/Lodestar and Hamlet systems further to provide smaller awarding bodies with adequate systems to undertake the activities listed in paragraph 4.1. Development of a centralised system for record keeping was the item given the highest priority by the industry bodies responding to the survey.

- 4.13 It would be helpful if these two systems were able to share data with the mainframe systems of major awarding bodies. Although joint certification has so far meant that two awarding bodies act together, there is also the possibility that in future there might be a 'vertical' division of labour between awarding bodies in that a small awarding body might deal with the 'live' candidates up to the point of printing and despatching awards, but that the records might be maintained and the statistics handled by the more powerful systems of a larger body. One of the advantages of this approach would be a greater guarantee that the records would be maintained for years to come, even if there were casualties among the smaller awarding bodies, e.g. for lack of economic viability. There is a precedent for this record handling in that when the old Joint Committees for National Certificates and Diplomas were abolished in the 1970s their record-keeping was taken over by City and Guilds.

#### **Possible markets**

- 4.14 The main market for standardised computerised systems are small awarding bodies. There are two main categories:
- Lead Bodies which become awarding bodies, but often without experience of what that entails.
  - professional bodies; although the larger ones have computerised systems there are many others which would need to have a system to handle examinations and awards within the NCVQ framework as well as dealing with membership issues.

#### **Possible follow-up action**

- 4.15 There are three possible activities which would be worth further consideration:
- a) Making funding available for existing proprietary systems for further development.
  - b) Providing financial assistance to small awarding bodies to obtain one of the software systems, including customisation of existing systems for them. One problem noted in the NQS feasibility study was that one small body not only hoped to have funding for software, but also for hardware and a computer operator. There are clearly issues about the viability of small awarding bodies which are much wider than computer assisted assessment itself.
  - c) At present few of the smaller bodies seem to be fully aware of the potential, or even the existence, of the NCVQ/Lodestar and Hamlet systems. Better dissemination of information about these systems might ensure that the more viable bodies purchased one of them rather than developing a one-off system of their own, which might be more difficult to link in future into any network of awarding body computer systems which might emerge.
- 4.16 Project specifications have not been prepared for paragraph 4.5 (a) and (b), because the issues require further consideration. Reference is made elsewhere in this report to the need for better dissemination of information.

## SECTION 5: LOCAL RECORDING AND TRACKING OF ASSESSMENT

### Introduction

- 5.1 Computers can assist in the management of the assessment process by the local centre by providing systems for the recording and tracking of candidate progress on individual units and elements. A subsidiary area of interest is in the use of computerised systems, including smart card technology, to record and monitor achievement of candidate's Individual Action Plans. These two aspects, although related, are considered separately in this report, as is discussion of systems designed for use by the candidate (paragraphs 5.42-5.46).

### Recording achievement of elements - main features

- 5.2 Computerised systems for recording and monitoring student/trainee progress on individual units and elements provide a means by which the local centre can keep track of its students. The system needs to hold information about the individual student (name, date of birth, address etc.) and may also hold information on the student's previous qualifications and perhaps work placement information. The system also holds information about the qualifications and units which the student is aiming to achieve, and the text of individual elements, sometimes including performance criteria and range statements. The local centre uses the system to record the date on which each student achieves each element within the target units.
- 5.3 Systems of this type allow for the production of reports on individual students, and also reports related to groups of students, showing which students still need to master which elements.
- 5.4 A number of the systems which exist have been developed to cope also with profiling, records of achievement and recording of success on BTEC courses, and therefore have options for recording student progress in terms of examination marks or assignment results.
- 5.5 Systems designed for use by colleges, rather than managing agents, may allow for student information to be imported from the college's Management Information System.

### Contribution to improving access, efficiency and effectiveness

- 5.6 Systems for recording candidate progress and achievement do not make any contribution to the **quality** of the assessment process. They do, however, contribute to the **efficiency** with which the system is administered. The survey of users identified the volume of paperwork as a major problem within the NVQ system, and discussions with individual colleges have confirmed that recording and tracking candidate performance is a major concern.
- 5.7 Provision of computer-aided recording systems also makes a contribution to saving assessor time, another problem area identified in the survey.

#### Recording achievement of elements - current availability and use

- 5.8 A number of proprietary systems are available to enable local centres to carry out this recording and monitoring. Examples are Rocket and Standards. A report from the Further Education Unit, *A Brief Guide to Issues in Establishing Computer Aided Records of Achievement*, published in late 1990, provides brief details of some of these systems, together with criteria for selecting a system and points to consider for colleges planning to develop their own systems.
- 5.9 The systems identified are intended to work on standard IBM compatible PCs; many are based on standard database packages, such as dBase or Dataease, but are 'customised' for use in recording candidate achievement.
- 5.10 In order to operate these systems, the centre needs to have available in a suitable electronic format, the IDs and titles of units with which it is dealing and the numbers and text of the elements. These may be typed in at the centre, but the more usual procedure is for the software producer to undertake the inputting of the text and to supply the centres with copies of the files which they need. The Hairdressing Training Board markets what appears to be a similar system (not seen), customised for Hairdressing NVQs.
- 5.11 The survey of users identified that 45 centres were currently using systems of this sort, the most frequently mentioned being Rocket, and a further 11 were considering their adoption.
- 5.12 Outside the NVQ system, an LTU-funded project 'Hypermaths' (also called 'Hypersmile') enables teachers to track the progress of pupils in learning tasks for National Curriculum Mathematics. The system works on an Apple Macintosh and records each pupil's achievement on each of a 'bank' of learning resource materials.

#### Recording achievement of elements - on-going research

- 5.13 No formal on-going research into this area is known. It may be expected that the software producers will be alert to the need to respond to the demands of the market.
- 5.14 Candidate recording systems will need to respond to any change in awarding bodies' requirements concerning the personal data collected about each candidate. The Employment Department is embarking on a National Vocational Qualifications Information Systems (NVQIS), collecting information from awarding bodies about qualifications awarded, and one of the medium-term possibilities for the development of this Survey is to request awarding bodies to collect more information about individual candidates, for example by ethnic origin, disability, mode of study and type of funding.

#### Recording achievement of elements - potential for further development

5.15 It is clear that the demand from centres has resulted in the production of a number of packages, and that there is no need for Employment Department financing for further development. There are, however, a number of respects in which initiatives by the Employment Department could contribute to a streamlining of the system:

- Awarding bodies could be encouraged to accept printout from these recording systems as part of the log book record for individual candidates, rather than insisting on manual completion of the awarding bodies' own log book.
- The system would be streamlined and potential errors reduced if the text of unit titles and elements were supplied on floppy disc to participating centres, either by the awarding body or by the Lead Body.
- Alternatively, a mechanism might be found by which the text of the elements could be downloaded from the NCVQ database, but this would require clearance of copyright issues; there is also sometimes a time lag between the announcement of a new qualification and its availability on the NCVQ database.
- As identified in the FEU report, there is almost certainly a need for agreement on consistent data formats between the awarding bodies, Lead Bodies, colleges and software system suppliers.
- The format for recording student details also needs to be consistent with DES standards for the student records, which are supplied to them as part of the Further Education Statistical Return.

5.16 It is becoming increasingly apparent that recording of candidate achievement needs to take account of range statements as well as the content of the elements. Software systems at present do not make provision for this, and it might be desirable for the Employment Department and NCVQ to initiate discussions between interested parties as to the necessary format for recording information about range covered, both in paper-based and computerised recording systems. Several centres participating in the survey suggested that a standard format for logbooks would be helpful.

5.17 Existing candidate recording software is used only to record assessment decisions made. As assessments are increasingly required to cope with ranges as well as elements, there may be a need for guidance to assessors on the sufficiency of evidence. This is covered in Section 11 of this report.

#### Recording achievement of elements - possible markets

5.18 Almost any centre is likely to be interested in computerised systems for recording candidate achievement; this was the item which attracted most support from centres completing the questionnaire. Development is likely to be highly cost effective because generic systems can be produced, with the text of the elements for each NVQ read in from text files.

#### Recording achievement of elements - possible follow-up action

- 5.19 The Employment Department could contribute to the development of this aspect of computer assisted assessment by disseminating information about what is available. The FEU report is not being published in full, although it is available on request and is to be mentioned in an FEU Newsletter. It is clear that many colleges are unaware of what is available and it would be helpful to produce a small directory. This could, however, form part of the proposed wider directory of CAA-related products.
- 5.20 Outline specifications have been prepared for follow-up projects, as follows:
- A coordinating and data standards project to make the development and use of recording systems more efficient (G5.1).
  - Investigation of the place of range statements in log books (G5.2).
  - A directory of candidate recording software (G5.3).

#### Individual Action Plans - Main features

- 5.21 There are systems in existence to record, in electronic format, titles and/or IDs of the qualifications and units which the student/trainee is aiming to achieve and which form part of his/her Individual Action Plans. Systems of this type are being used mainly within the pilot schemes for training credits, and hence may deal also with the recording and monitoring of the expenditure of cash units by the trainee. (Units of finance are normally linked to units of delivery, which are not necessarily the same as units of accreditation).
- 5.22 Systems for recording and monitoring Individual Action Plans are very similar to those for recording achievement on individual elements, but Action Plan monitoring does not go down to element level. Action Plans may also include qualifications or units which are not NVQs and which are not divided into elements. Some qualifications, particularly older ones or those which are supplementary to the main qualification taken, may not be divided into units.

#### Contribution to access, efficiency and effectiveness

- 5.23 As with the systems previously described, software for monitoring Action Plans contributes to the efficiency of the system, rather than to the quality of assessment.
- 5.24 The SOLOTEC smart card system for recording candidate achievement is also intended to promote 'ownership' of training and qualifications because the trainee him/herself holds the smart card. Because of its novelty value, it can play a part in promoting the Training Credit philosophy, which is itself intended to increase the take-up of both training and qualifications. Other systems may be intended to facilitate flexible course provision, encouraging access to education by non-traditional students by providing early recognition of achievement.



### Individual Action Plans - current availability and use

- 5.25 In the University of Bologna, an experimental smart card system controlling examinations on demand for students with Individual Action Plans has been sufficiently successful for it to be extended to the full University (which has 80,000 students). The experimental system was included as a case study in the Employment Department (LTU) report *The Learning Credit Card*. At the Institute of Higher Education in Tashkent, the student can choose certain parameters (e.g. course level, frequency of assignments) and receive an individualised study programme.
- 5.26 The most advanced UK example of this type known to the project is the smart card system being introduced by SOLOTEC (South London Training and Enterprise Council) as part of its training credits pilot. The candidate's Action Plan is defined in terms of target qualifications and units, and the IDs of these qualifications/units recorded on the smart card. When the smart card is inserted into a reader attached to a PC with the necessary software, the target IDs can be read and decoded into qualification and unit names by interacting with a file derived from the NCVQ database. The candidate's achievement of each of the units is recorded on the smart card by staff at the centre.
- 5.27 In the present SOLOTEC system the individual candidate's Action Plan is drawn up with the use of a specially commissioned software module, which writes a small file for each candidate for transfer to the smart card software. However, NCVQ is developing a 'book mark' facility for its database, to allow the target units to be identified and tagged on the database, and then written to an Action Plans file for transfer to the smart card.
- 5.28 It is known that Action Plan software, not involving the use of smart cards, is being prepared for use in the training credits pilot by Bradford TEC and LEA.
- 5.29 There is interest from both Wirral and North East London Colleges in recording student achievement in terms of units or non-NVQ modules, encouraging flexible provision and contributing to positive feedback to students.

### Individual Action Plans - on-going research

- 5.30 SOLOTEC has provided a demonstration of its system to other TECs and it is quite likely that smart cards will feature in some TEC bids for the second round of Training Credits Pilots.
- 5.31 A feasibility study on the use of smart cards in ET is being conducted by Transcend Technology Ltd for the Employment Department. The report is due in early November 1991.
- 5.32 GES is involved as a partner, with French (main contractor) and Italian partners, in a submission to the FORCE programme of the European Commission on the use of smart cards for employed adults. If the proposal is accepted there would be two years of action research.

- 5.33 One possible further development of the National Vocational Qualifications Information System is the institution of a unique national identifier for all vocational qualifications. This would be primarily intended to facilitate national monitoring, enabling calculation of the number of **qualified people rather than of qualifications gained and identification of patterns** of achievement. However, it would also simplify some aspects of centre administration.

#### **Individual Action Plans - potential for further development**

- 5.34 It is clear that any development of software for Individual Action Plans should be made compatible with the systems described earlier for recording success on individual elements.
- 5.35 Problems encountered in the introduction of the SOLOTEC system have included:
- The fact that unit IDs and qualification IDs may be changed, for example if the ID which the awarding body or Lead Body proposes to use is not known to NCVQ at the time when the database record is set up.
  - Initial difficulties in identifying all the qualifications which may be required within the system.
  - The existence of older-type qualifications which are not divided into units.

The last problem is likely to be a transitional one, disappearing when the NVQ system is fully in place.

- 5.36 Both North East London and Wirral Metropolitan Colleges had also encountered difficulties in introducing their flexible learning provision because of the existence of older qualifications not divided into units. Wirral is considering dividing qualifications into 'soft modules', and North East London College had already undertaken similar work in certain subject areas as part of its Credit Accumulation Project (CAP). It is possible that there are other colleges throughout the country similarly engaged in dividing qualifications into modules using their own systems, but this is not seen as a useful long-term development.
- 5.37 As part of the SOLOTEC system, additional database records have been added to the NCVQ database for use within the SOLOTEC area, and to allow interaction with the smart card. The new records have included those GCSE and A level subjects which are available at colleges within the SOLOTEC area, units for Employment Department approved qualifications of potential interest to SOLOTEC trainees, and some additional qualifications (including pre-vocational) which are of potential interest to special needs trainees or which cover subject areas not as yet available as NVQs.

- 5.38 There is likely to be growing interest from centres in linking academic and vocational qualifications within students' Action Plans. General NVQs and the proposed Ordinary and Advanced Diplomas will further increase this trend. Any extension of the SOLOTEC smart card system to other TECs would almost certainly require the addition of further qualifications and units to the associated database which might eventually include all GCSE and A/AS levels as well as GNVQs.

#### **Individual Action Plans - possible markets**

- 5.39 With the extension of training credits, both within Youth Training and within Employment Training, and increasing interest in flexible delivery systems, there is a large potential market for systems of this type.

#### **Individual Action Plans - possible follow-up action**

- 5.40 Much of the necessary action to develop these systems is already being taken, in some instances with funding from other branches of the Employment Department. However, a particularly useful initiative would be to encourage Lead Bodies and awarding bodies to declare the unit and qualification IDs which they propose to use when material is submitted to NCVQ for accreditation. This would avoid subsequent changes to records in the NCVQ database and linked systems.
- 5.41 Much necessary standardisation work has already been done, but the requirements of Individual Action Plans should be borne in mind in discussions on standardisation related to recording of success on elements, as described in paragraphs 5.15 and 5.17 above.
- 5.42 An area of interest mentioned by Wirral Metropolitan College was the linking of the Action Plan for the individual student to recommendations of relevant learning material. One source of relevant information is the 'Matching Programme' being undertaken by the Learning Systems and Access Branch of the Employment Department. This programme sponsors projects related to individual subject areas in which analysts with subject knowledge identify learning material, and match them for their relevance to the standards of competence in NVQs. An outline specification (G5.4) is given for a demonstrator project linking this information to a computerised Action Plan and guidance system in one or two subject areas.

#### **Systems for use by candidates**

- 5.43 The systems described earlier in this section are intended mainly to enhance the management of the assessment process by the local centre. Related systems could, however, be used to give the candidate greater autonomy in managing his/her progress.

- 5.44 There is evidence (for example from North East London and Wirral Colleges) of interest in the use of the DES National Record of Achievement within colleges, and this may be taken up at national level in the future. The DES National Record of Achievement calls for free text entries by the student on pages of the record. It was suggested at North East London College that it would be valuable to have a software program to enable students to draft and check their entries, and then to produce them in the appropriate format for printing on to the sheet. The ability to produce a well-presented, error-free copy was seen as a powerful motivating force. However, as this idea is not firmly linked to NVQs, no detailed project proposal has been prepared.
- 5.45 A system which combined candidate recording software (paragraphs 5.2 to 5.20) with guidance for the Accreditation of Prior Learning (Section 12) could be used to help candidates to manage the assembly of their own portfolio of evidence, guiding them to areas where additional evidence is needed and suggesting suitable types of evidence. Encouraging candidates to assume responsibility for the process of evidence collection would also improve motivation and could reduce the work required of the assessor. This is seen as possibly longer-term development and no outline project specifications has therefore been prepared.

## SECTION 6: TEST DESIGN AND PRODUCTION

### Main features

- 6.1 Two main categories of product are considered in this section: computer-assisted test compilation and computer-assisted item generation. Computers are also used in more conventional ways to produce assessment material by word processing or desk-top publishing. However, these uses are well established, would not require Employment Department funding, and are therefore not considered in this report.
- 6.2 Computer-assisted test compilation takes existing questions from a 'bank', and combines them according to pre-set rules to produce a new combination of questions. This may be as part of a system for computer delivery of assessments (as in Question Mark), but a similar technique can be used to compile tests for printing and conventional delivery. Although the technique has been used almost exclusively for objective questions up to now, it could equally be used for more conventional ('constructed answer') questions, or even for assignments.
- 6.3 In a computer-assisted test compilation system, each question, or assessment task, is written in advance and any necessary validation takes place before the question is entered on to the computer. The compilation system should provide for rules (sometimes referred to as a test specification) to determine the way in which questions may be selected. This will most usually be in terms of sections of the syllabus or topics, providing (for example) that there will be two questions in each test on a given topic. The system may provide for statistical analysis about each question to be held and used to calculate the predicted mean mark and standard deviation of the test.
- 6.4 Computerised item generation systems can be used to generate new questions or assessment material following a set of pre-determined rules. The questions or tasks generated in this way may then be compiled into tests either manually or using the techniques for computerised test compilation. By means of computerised item generation a number of parallel versions of each question or task can be produced as required.
- 6.5 Computerised item generation could be used either for computer delivery of assessments or to produce questions or tests for printing and conventional delivery.
- 6.6 Areas in which computerised item generation would be viable include:
- numerical questions
  - production of text for tests of keyboarding or comprehension
  - audio delivery of verbal material for similar purposes
  - alternative scenarios in crisis management assessment
  - production of multiple-choice or other objective questions in accordance with pre-set rules.

- 6.7 It should be noted that in each case the 'rules' must be determined in advance by the designers of the system, and that the components from which the questions or tasks are to be generated must be determined by subject experts.

#### Contribution to improving access, efficiency and effectiveness

- 6.8 For the awarding or assessing body wishing to produce large numbers of parallel tests for conventional administration, computerised item generation offers improvements in **efficiency**.
- 6.9 However, the most significant contribution of these techniques would be in conjunction with local, on-demand assessment procedures to reduce the perceived problems of security and likelihood of cheating. If there are large numbers of parallel versions of any given test, it becomes less likely that the candidates will gain an unfair advantage by having seen the test material before (either illicitly or because they have attempted the test previously). On knowledge tests, it can be argued that if the candidate knows the entire content of the question bank then he or she deserves to pass anyway. It follows that item and test generation techniques, when combined with computer delivery of assessment or other on-demand assessment systems, can improve **access** to assessments without reducing their effectiveness.
- 6.10 Computerised test compilation and item generation techniques may also help to improve the **effectiveness** and fairness of assessment by ensuring that all parallel versions of the test are set to the same rules, hence reducing variability in the standard and coverage of tests set on different occasions. There is, however, some doubt as to whether multiple-choice questions based on a given set of rules and content will necessarily all be comparable in difficulty. For example, a question about the properties of a little-used steel alloy would be more difficult than an apparently comparable question on the properties of low carbon steel.

#### Current availability and use

- 6.11 One of the most significant existing examples of computerised test compilation in the UK is the City and Guilds CAIBAR system (Computer Assisted Item Banking and Retrieval). This is used to produce City and Guilds' multiple-choice tests from their banks of items, the tests being printed for conventional administration. The system provides for a 'test specification' which defines the number of questions to be set on each topic, and stores information about question analysis from which the predicted mean and standard deviation of the compiled test can be calculated.
- 6.12 At the other end of the scale, Question Mark can select individual items from a sub-bank for computer administration, but its techniques appear to be less rigorous than those of CAIBAR, and there is no provision for taking account of the statistical analysis of the questions. City and Guilds' Inside Information provides simple selection of questions from a bank in a subject-specific context.
- 6.13 Several systems available in the USA and not examined in the project can store item banks and compile tests from a pre-set 'strategy template'. Reference is made in Section 9 of this report to those systems which also deliver tests at the computer.

- 6.14 Use of item generation techniques to produce assessment material is much less well-developed, although some of the techniques have been used for other purposes. As far as is known, the only actual use of item generation techniques is in connection with numerical questions or with other parameters which can be expressed in numerical terms. The Micropat pilot test selection suite developed by David Bartram for the Ministry of Defence includes several sub-tests in which the individual questions can be generated through the random number generator feature. It is understood that the Mathematical Association of America has a system with similar but more extensive functions.
- 6.15 Psychometric Research and Development (PRD) have experimented with software to generate unpredictable sequences of events (appropriate, for example, for generating crisis management scenarios), and also new passages of text following given rules.
- 6.16 As far as is known, there are no software systems to produce multiple-choice or other objective tests from given rules, and this is an area which deserves exploration. City and Guilds multiple-choice question writers in the early 1970s had identified that there were certain categories of question which could be produced in a number of different variants with slightly differing content; for example, candidates could be presented with diagrams showing four out of a range of commonly used symbols for electrical components, and asked to identify the one which represented a resistor; a similar question could ask for the identification of a capacitor symbol, or a transformer.
- 6.17 There may be some limited scope for using the same techniques to produce variants of non-objective questions, particularly where the questions are partly numerical. Experienced examiners have in fact been doing this for years.

#### **On-going research**

- 6.18 On-going work appears to be very limited in extent, at least as regards the actual use of these techniques for assessment purposes. PRD is working on an interactive video to teach and assess crisis management techniques in a chemical plant.

#### **Potential for further development**

- 6.19 There is potential for further development of the test compilation software to combine the more rigorous characteristics of CAIBAR with an assessment delivery mechanism such as Question Mark. It is recommended that the Employment Department should consider commissioning a 'shell system' with all the necessary features for item banking, test compilation and delivery, feedback of candidate responses to the awarding body and item analysis. This might incorporate elements of existing software and producers of existing relevant software could be amongst those invited to tender for this work. Software production should, however, be preceded by consultation with potentially interested parties to develop a more detailed specification.

- 6.20 Shell software could also be developed to enable question writers to generate different parallel objective questions from sets of related information. However, there is a need for research into the extent to which techniques of this sort would produce questions of comparable difficulty, and this would be best undertaken in conjunction with an awarding body which has some experience of setting multiple-choice questions and has access to a fairly large pool of potential candidates for trialling of alternative question forms.
- 6.21 Questions which have a numerical basis are the easiest to adapt for automatic item generation, and are more likely to be comparable in difficulty. The proposed shell could be extended to incorporate random number generation for both objective and conventional questions either for computer delivery of the assessment, or for printing for conventional delivery. Features should include the possibility of including simple graphics in which the parameters would be determined randomly within pre-set rules (for example drawing a circle with a radius of any whole number within a given range).
- 6.22 Techniques for generating alternative but parallel versions of text should be investigated as a means of producing test material for:
- typewriting tests (including basic keyboarding, lower level typewriting, word processing and, with the addition of audio output, audio typing)
  - comprehension of either English or a foreign language
  - text from which an accounting candidate might be asked to draw up a balance sheet or (at a lower level) to fill in a petty cash book.
- 6.23 Audio output of computer-generated text could be investigated as a source of test material and delivery for oral comprehension, audio typewriting, and also spelling in either English or a foreign language.
- 6.24 Computer generation of text to describe a scenario could be used to generate stimulus material of the 'case history' type to be used in relation to conventionally delivered tests in areas such as Caring. Generation of scenarios for use in areas such as crisis management could form part of any work undertaken in that area, (see Section 10)

#### **Possible markets**

- 6.25 Although the obvious potential market for these techniques is amongst awarding bodies who wish to produce externally set tests, the techniques could also be used as a way of producing assessment resource banks to be sold or licensed to centres or TECs. The survey of users confirmed that centres would welcome provision of resource banks of assessment material.
- 6.26 Those techniques which relate to knowledge assessment or numerical skills can be used in almost any subject area. Other techniques relate more naturally to secretarial and office qualifications or to communication and language work, but additional areas where such techniques could be useful may become apparent as the techniques are developed.



## Follow-up action

6.27 It is recommended that the Employment Department should consider a series of follow-up projects, as discussed above (paragraphs 6.19-6.26), but linking in some cases to projects for computer delivery of assessment. Some of the follow-up work would be dependent on the results of the initial projects. Outline specifications have been provided for the following initial projects:

- production of specification and 'shell' software for objective question writing/editing, test compilation, computer-delivery, scoring and analysis (G 6.1)
- item generation by rules (G 6.2)
- text generation for typing and English comprehension (G 6.3)
- text generation for accountancy problems (G 6.4)

## SECTION 7: COMPUTER DELIVERY OF ASSESSMENT - GENERAL

### Introduction

- 7.1 One of the major potential roles for computer assisted assessment within National Vocational Qualifications is in the delivery of assessment at the computer.
- 7.2 Because there are so many potential developments to be discussed under this heading, the topic has been divided into three separate sections, covering computer delivery of skills assessment (Section 8), knowledge assessment (Section 9), and simulation (Section 10). This section sets out some of the principles governing computer delivery of assessment, and discusses aspects which are common to all three categories.
- 7.3 The major distinguishing feature of computer delivered assessment is that the assessment tasks are presented on screen (although supplementary written material may sometimes be necessary), and that the candidate responds at the keyboard. By contrast, the aids to test production discussed in Section 6 may result in the production of camera-ready copy for test printing and conventional test delivery.
- 7.4 The main benefits from computer delivered assessment are gained when the marking and reporting of results can be undertaken by computer. This means that the candidate's response must either be by selection of one or more options from a given list, or by inputting at the keyboard one or more characters which can be marked objectively. It is possible for the computer to mark numerical input, keyboarding accuracy and some other categories of response which can be defined precisely. Input of an answer which is a single word or phrase may also be computer marked if an exhaustive list of acceptable answers can be input, but this can cause difficulties. The technique can be used more readily in computer-based training, where computer failure to recognise a valid response by the user may be extremely irritating but is not as critical as it would be in assessment for a qualification.
- 7.5 In circumstances where computer marking is not feasible, the computer may still have a role in the storage and delivery of assessment material, but this needs to be combined with discussion and follow-up questions with an assessor or by examiner marking of written responses.
- 7.6 Computer delivered assessment may be adaptive, in that the candidate's response to the first questions or tasks presented will determine what questions or tasks are presented next. Adaptive techniques can be used for tests of knowledge or of abilities, and computer simulations are also 'adaptive' in that the program branches according to the candidate's responses.

## Contribution to improving access, efficiency and effectiveness

- 7.7 Provision of computer delivered assessment could lead to an improvement in **effectiveness** by improving the quality and consistency of assessment. At present it is likely that colleges, managing agents and specialist training providers and even, in some instances, trainers within firms, devise their own assessment tasks in order to ensure that assessment covers all the elements and performance criteria within the unit. While many of these assessment tasks are no doubt well planned, well presented and closely related to the needs of the workplace, it is likely that a proportion are unclear, too easy or difficult or out of line with the intentions of the standards of competence.
- 7.8 Production of nationally available computer-delivered assessments could be used as a way of disseminating **best practice**, drawing on the expertise of those who are ~~most experienced in the~~ relevant occupational area. Such assessments would also provide a 'bench mark' against which assessors and verifiers could judge the appropriateness of centre-devised assessment tasks or of naturally occurring workplace activities. This should also help to ensure that the type of activity required of candidates for the qualification is more consistent across all centres.
- 7.9 Knowledge tests and simulations, whether computer delivered or traditional, can provide a way of extending the **range** covered by the whole assessment process. The candidate can demonstrate competence in one application in the workplace and then demonstrate the ability to apply his or her knowledge skills to other situations in computer delivered assessment.
- 7.10 A specific way in which computer delivered assessment can extend the range of assessment is in relation to **circumstances which occur only rarely** in practice, or which involve **crises or dangerous situations**. Simulations, in particular, can contribute to the assessment in this area, and the surveys confirmed that assessments of this type caused difficulties for centres.
- 7.11 The assessment of **underpinning knowledge and understanding** has been an emerging issue in the recent development of NVQs, and will be of increasing importance as the NVQ Framework is extended to give greater coverage at the higher levels of competence. Assessment of underpinning knowledge is one means by which the range of applications covered by the assessment can be expanded without a great increase in cost and assessor time. In some NVQs there is also likely to be a need to test understanding and the ability to apply knowledge to new situations which are not readily apparent from workplace observation.

- 7.12 Provision of an assessment resource bank of computer delivered knowledge tests may also be a means of improving the **quality** of assessment of underpinning knowledge by contrast with locally devised assessment. Although locally devised knowledge assessment can be more responsive to the candidate's individual circumstances, it may well be of lower quality and less consistent than centrally devised assessment material, which can draw on a wider range of subject expertise. At present, much assessment of underpinning knowledge in NVQs is by means of internally devised oral questioning, and oral tests are notoriously difficult to standardise. However, it would be necessary to ensure that any computer delivered assessments of knowledge do not unnecessarily penalise candidates with a low level of literacy where this is irrelevant to the knowledge or skills being assessed.
- 7.13 Another important advantage of introducing computer delivered assessments would be the gains in **efficiency** resulting from the reduction of assessor time and the cost of assessment. Although it is unlikely that computer assessment will be able to replace workplace observation completely, other than in a very few units, there will be some units where assessor time can be reduced to a minimum, with a consequent saving in costs. Computer delivered assessment could therefore help to solve two of the three biggest problems in NVQ assessment reported by centres.
- 7.14 It should also be possible to achieve some reduction in the time which the candidates spend on assessment; specially designed computer-delivered assessment can cover a wider range of situations than would occur naturally in the workplace over a limited period of time.
- 7.15 In units where the majority of the assessment could be by computer, it might be sufficient for this to be supplemented by signed statements from employers or workplace supervisors who were not specially trained in assessment techniques. Such statements might, for example, confirm that the candidate regularly observed safety, security and backup procedures in connection with keyboarding. Combined with a computer delivered test they might cover the whole requirements of a keyboarding element or unit. A reduction in the reliance on trained workplace assessors would have a significant effect on improving access to NVQs, particularly for candidates in small firms and those who are not in employment. The availability of test material at open assessment centres (e.g. in colleges or TEC offices) would also improve opportunities for access for this category of candidates.
- 7.16 Introduction of computer delivered assessment would also make it easier to ensure that assessment could be available **on demand**, whenever the candidate is ready, which would also facilitate candidate access to qualifications. Immediate feedback of results and the opportunity to retake the test when required would further enhance access and flexibility. Without computer delivered assessment, a move to make assessment more rigorous might lead to a greater reliance on infrequent fixed date externally set examinations, which impose constraints on the flexibility of learning.

- 7.17 Ideally, computer-delivered assessment would be made available, not only for candidates following courses but also in 'open assessment centres'. These might be based at colleges or TEC offices and would provide assessment facilities (including APL) for all candidates irrespective of their mode of learning or the facilities provided by their employer.

#### **Current availability and use**

- 7.18 The current availability and use of computer delivered assessment is discussed in more detail in the next three sections, dealing separately with skills assessment, knowledge assessment and simulations. In general, however, there is very little computer delivered assessment for certification purposes in the UK. Some delivery systems are available for the administration of tests at the computer, but most of these are specific to a small number of applications, and many are used only in tests of general ability. There are, however, a great many techniques used in computer-based training which have the potential to be developed for computer delivered assessment.

#### **On-going research**

- 7.19 On-going research is described in more detail in Sections 8-10 but, again, there is at present only limited activity specifically related to the use of computer delivered assessment for purposes of certification.
- 7.20 One general initiative is that Working Knowledge Transfer Ltd is involved in the early stages of a project to identify aspects of the Management Charter Initiative competences which could be assessed by computer. Wirral Metropolitan College is applying for TEED funding to support a similar project covering a wider range of subject areas. The Royal Institute of Public Health and Hygiene is exploring the potential of computer delivered knowledge tests and joint packages for teaching and assessment.

#### **Potential for further development**

- 7.21 In all three major aspects of computer delivered assessment there is very considerable potential for further development, drawing upon the lessons to be learned from relevant applications in other areas, particularly CBT and general ability testing. Recent Employment Department funding in relation to technology-based training has tended to concentrate on the use of more advanced technologies, such as interactive video and simulations, and there are lessons to be learned from these developments for computer delivered assessment. However, there is also a very great potential for the use of lower level technology which can address a significant proportion of the needs for computer delivered assessment at lower cost. Some possible developments might ultimately be self-funding because of the savings they produced in (for example) assessor costs.

- 7.22 There are, however, other areas of potential CAA development where interactive video or high quality computer graphics are required and where the cost of developing a package solely for assessment purposes would be excessive. In those areas, consideration should be given to developing combined packages, intended for both learning and assessment purposes. Consideration should also be given to developing 'add on' assessment modules which make use of existing video and graphics material. Where existing learning material has a significant assessment element, it might be 'licensed' by the relevant awarding or Lead Body who would state the extent to which successful completion of the material would be regarded as evidence of competence in relation to specific elements.
- 7.23 A useful initial development would be for the Employment Department and NCVQ to work with awarding and Lead Bodies to define Acceptability Criteria for the use of learning material or combined material for assessment purposes. These would be refined once the bodies had actual experience of 'licensing' specific packages.
- 7.24 Any developments in using material for both learning and assessment need to take account of the concerns expressed by the Learning Technologies Unit of the Employment Department who have suggested that this could distort the educational or training value of the material. There would be a danger that packages would be devised with the narrow requirements of the assessment in mind rather than with broader educational aims and the 'learner centred approach' might be lost. If this problem could be overcome, candidates who had successfully completed a learning package would undoubtedly welcome the opportunity to use this as evidence for NVQ certification.
- 7.25 At a much simpler level, computer technology could be used to supply banks of assignment material to centres on floppy disc or CD-ROM. The assignments could be used as supplied or could be 'customised' by the centre to meet local needs.

#### Possible markets

- 7.26 For ready-to-use assessments, packaged and sold in the form of assessment resources, the possible markets are colleges, YT/ET providers, individual firms, and indeed any organisation which is a centre for the assessment of candidates for NVQs. The user survey established that 75% of the centres who responded had IBM compatible computers which could be made available for use in assessment, and 72% of respondents would welcome the provision of computerised tests of underpinning knowledge. The availability of more specialised equipment was much lower; 15% of centres had interactive video equipment and only 11% had a CD-ROM player.
- 7.27 In addition to the use of computer delivered assessment for their own candidates, colleges and TECs might be interested in setting up open assessment centres as a service to local business or to provide facilities for candidates who might not have ready access to assessment within their own organisation. Clearly this development would only take place if there were a sufficient volume of assessment material available for use, although this could include paper-based material as well as computer delivered assessment.

- 7.28 Awarding bodies and Lead Bodies are a potential market for computer delivered assessment, either as a means of distributing their own externally set assessments or as partners in the provision of an assessment resource bank. In either case, any development of computer delivered assessment material needs to be undertaken in conjunction with relevant awarding and Lead Bodies so as to ensure that the materials produced are acceptable as 'evidence' of competence for their qualifications.
- 7.29 In general, there was a significant level of interest in the potential for computer delivered assessment in the Industry Bodies, and individual Bodies might well wish to play a major role in the development of computer delivered assessment for applications of particular interest in their own subject areas.
- 7.30 The market for computer delivered assessment of underpinning knowledge would spread across all subject and occupational areas, although perhaps of least relevance at Level 1 in the NVQ Framework. Assessment of skills at the computer is of greatest relevance in the areas of Business Administration and Information Technology, but other occupational areas may also require limited amounts of assessment in these subject areas. For example, some of the skills required in Business Administration are also required in Retailing. Computer delivered simulations are of benefit in a smaller range of specialist situations involving infrequent occurrences or high risk, but also in tasks such as fault-finding, where the assessment process on computer can be made adaptive.

#### Possible follow-up action

- 7.31 Sections 8-10 of this report, and related outline project specifications contain proposals for follow-up action which are specific to the three main categories of computer delivered assessment (skills, knowledge and simulation). There are, however, a number of developments which relate to all computer delivered assessment and which it would be helpful for the Employment Department to promote (in association with NCVQ as appropriate), and these are described briefly in this section.
- 7.32 Sections 8-10 contain a number of individual proposals for follow-up action which are relevant to NVQs in Business Administration at Levels 1 and 2. It would be particularly valuable to provide computer delivered assessment material in this occupational area because it is one in which candidate numbers are large and in which there are relatively high percentages of candidates who are not employed but are following courses at colleges or training centres. For these candidates, assessment in the workplace may be difficult to arrange or expensive; often it involves the tutor in going out to observe the candidate in the workplace during work experience. It is therefore suggested that the Employment Department should initiate a project to define, with the aid of subject experts and representatives from the awarding bodies, those units and elements within the Business Administration NVQs which could be assessed by computer delivered assessment. The project team could then draw on the work of some of the other proposed projects to combine the products developed by them into an assessment resource covering as much as possible of the Business Administration NVQs, for distribution possibly on CD-ROM.

- 7.33 Drawing on the lessons to be learned from the Business Administration project and on the information contained in this report, other projects could be launched to identify the potential for computer assisted assessment in other specific NVQs. A particularly useful area to cover would be Retailing, where the evidence of the centre survey and several items of anecdotal evidence suggest that assessment in the workplace is particularly difficult and costly to arrange.
- 7.34 It would clearly be desirable that all developments of computer delivered assessment for accreditation purposes should follow, as far as possible, similar formats and procedures. For example, the development of computer delivered assessment should always be accompanied by the development and availability of practice material to allow the candidates to familiarise themselves with the software and necessary procedures before they undertake an actual assessment. An early priority should be for the Employment Department and NCVQ to discuss with awarding bodies and centres a Code of Practice to be observed by producers of computer delivered assessment; it is likely, however, that this will require refinement and extension as more experience is gained of computer delivered assessment.
- 7.35 Security will be a serious concern for awarding bodies, and must be given due weight if the credibility of awards is to be maintained. One mechanism for ensuring security is to establish a Code of Practice for assessment centres; adherence to this Code of Practice would be a condition for approval of the centre to offer computer delivered assessment. It would be desirable for all awarding bodies in the NVQ system to work to the same Code of Practice and to apply a common approval mechanism for centres wishing to provide computer delivered assessment. Indeed, a centre should need to apply only once for approval, and this approval should be valid for all awarding bodies.
- 7.36 It would also be helpful to discuss with awarding bodies, TECs and centres the security and infrastructure requirements which are required for any large scale implementation of computer delivered assessment and to reach an agreed statement.
- 7.37 If the concept of an open assessment centre, both for computer delivered assessment and perhaps for more traditional assessments and APL, is felt to be helpful, it would be beneficial to provide initial funding for a small number of pilot assessment centres. It would, of course, be necessary to time the launch of such centres in such a way as to ensure that there was sufficient assessment material available for them to offer a reasonable range of services.
- 7.38 In the longer term, and drawing on the lessons learned from the pilot centres, the Employment Department might wish to consider development funding for other colleges and TECs to help them to set up open assessment centres, and in particular to enable them to acquire the necessary equipment. No detailed proposal is given for this as it would depend on the results of earlier work.



- 7.39 Subject to the reservations expressed above (7.24), the Employment Department should also consider collaborative projects in which learning material and assessment material were designed together. This would require collaboration with the relevant personnel in the Learning Technologies Unit, who have considerable experience of funding hi-tech applications. A preliminary stage would be to define the Acceptability Criteria for joint learning and assessment packages.
- 7.40 All specific developments of computer delivered assessment material, as detailed in Sections 8-10 of this report, should be designed so that they can accept both fixed assessment tasks and also assessment tasks generated by the techniques described in Section 6. The longer term aim would thus be to produce suites of assessment devices in which there were numerous versions of each assessment task.
- 7.41 At an early stage of the follow-up developments it would be valuable for the Employment Department to hold one or more dissemination workshops to make awarding and Lead Bodies aware of the potential for CAA and of specific on-going developments. This would also provide a forum for discussion of issues. Continuing dissemination activities would be required as more CAA products were developed.
- 7.42 Outline project specifications have been prepared for the following projects.
- An assessment resource for Business Administration (G7.1)
  - Analysis of the potential for computer delivered assessment in Retailing (G7.2)
  - Development of a Code of Practice for developers of computer delivered assessments (G7.3), statement of infrastructure and security requirements (G7.4), Code of Practice for centres (G7.5) and Acceptability Criteria for combined learning and assessment material (G7.6)
  - Dissemination workshop (G7.7).

## SECTION 8: DELIVERY OF SKILLS ASSESSMENT

### Main features

- 8.1 There are a number of skills which may naturally be exercised by use of a computer (keyboard proficiency itself being an obvious example), or which can be exercised at the computer with relatively little departure from reality. Examples of the latter include numerical and communication skills. There is considerable potential for such skills to be assessed at the computer with the marking performed automatically.
- 8.2 The main categories of skills considered under this heading are:
- secretarial skills, including basic keyboarding, typewriting and audio typing
  - skills in using computer software, including DOS, spreadsheets, databases, word processing, computer-assisted design and computerised accounting
  - simple computer programming
  - core skills, specifically numeracy, communication skills, information skills and problem-solving
  - skills related to accounts work
  - the underpinning skills related to other business activities, such as filing and mail handling.
- 8.3 There is no clear dividing line between skills assessment at the computer, and simulation (where the exercise of the skills does not normally take place at the computer). For convenience, this report is classing as 'simulation' activities where the computer responds to the candidate in a different way according to the candidate's initial choices. Under the heading of Simulation (Section 10) the report therefore considers assessment of fault-finding, plant operation and decision-making.
- 8.4 The skills discussed in this section can be exercised at the computer keyboard with very little departure from normal practice, and the assessment tasks can be designed so as to produce responses which can be marked objectively by the computer.

### Contribution to improving access, efficiency and effectiveness

- 8.5 The contributions of computer delivery of assessment to improving access, efficiency and effectiveness were discussed in Section 7. For computer delivered skills assessments the main benefits lie in the **reduction of costs and assessor time**, with some likely improvements in **quality and access**.

## Current availability and use

- 8.6 Many of the appropriate techniques and types of computer software have already been partially developed for use in computer-based training but use of these techniques for the purposes of assessment for accreditation is much more restricted. City and Guilds have a keyboarding test for which an electronic version is available, but it is still in the pilot stage. Tests for general and medical secretaries, developed for the employment agency Manpower assess word processing skills, but require marking by a human assessor. The most developed system is the Aequitas series of tests, produced in the United States by the Educational Testing Service and Kee software. Each of the Aequitas tests is specific to one proprietary software package, for example Word Perfect or Lotus 1-2-3, and tests the proficiency of the candidate in using the software, including associated keyboarding and communication skills. The exercise is computer marked and the resulting report indicates the software functions (such as centering and justification) which the candidate has mastered.
- 8.7 Computerised tests of keyboarding skills, proof-reading and filing-related skills are known to form part of two computerised test batteries, the Minnesota Clerical Assessment Battery and Genesys.
- 8.8 There are several batteries of general ability tests which include numerical, 'verbal reasoning' and other tasks with a strong similarity to the core skills of number and communication. However, these tests are used for selection, not for accreditation, are norm-referenced, and many of them are designed for graduate-level applicants.
- 8.9 There are a large number of CBT packages providing instruction in basic keyboarding, all of which provide exercises for the testing and marking of keyboarding speed and accuracy.
- 8.10 A number of CBT packages teach the use of specific types of software, including database packages, spreadsheets and word processing; an example is 'Teach Yourself Lotus 1-2-3'. These packages provide a simulation of the screen as it appears to the user in the ordinary user version of the software, instruct the user how to enter the necessary commands to perform specific functions, and warn if the command is incorrect. The examples seen do not have a strong assessment element.
- 8.11 Existing CBT packages also teach accountancy-related skills, for example teaching the user how to enter figures into a balance sheet and checking that the entry has been made accurately and in the correct column. Because the exercise simulates a paper-based recording sheet, it is not dependent on the use of specialist keys to perform particular functions, as is the case with database or spreadsheet use.
- 8.12 At least one CBT package (Filing: the Business Program) teaches and tests skills related to filing, for example arranging lists in alphabetical order.
- 8.13 Some existing CBT packages cover the core skills of numeracy and communication.

## On-going research

- 8.14 One current development in this area is that an adaptive test of numeracy skills is being developed by City and Guilds, using assessment material from its existing banks. The objective is to provide a computer-delivered test in which the questions presented to the candidate will be of a level of difficulty appropriate to his/her ability as judged by the responses to the earliest questions presented. The candidate can be routed to an appropriate group of questions to enable him or her to qualify for one of several levels of certificate in Numeracy.
- 8.15 A project being funded by the Employment Department and supported by ALBSU (the Adult Literacy and Basic Skills Unit) is developing teaching material for adult literacy students which makes use of CD-ROM XA and audio, as well as visual input. The product is being developed by Cambridge Training and Development. Other work being undertaken in adult literacy teaching and reported by ALBSU concerns the use of audio tapes to provide stimulus material for the assessment of oral comprehension, and this is also being used for business-related applications such as the development of skills in taking telephone messages.

## Potential for further development

- 8.16 There is considerable potential for the development of computer delivered skills assessment relevant to some occupational areas. Many of the software techniques are well established for use in computer-based training. The occupational areas where skills assessment could be computer delivered include some, such as Business Administration, in which candidate numbers are relatively large.
- 8.17 A test of basic keyboarding could be developed for use in assessment for Business Administration NVQs and others where keyboarding skills are needed. This might include Information Technology applications and Hotel Receptionists, and possibly also some staff in Retailing.
- 8.18 It would be relatively easy to develop computer delivered assessment of simple audio typing where all that was being assessed was the correct transcription of recorded speech.
- 8.19 Programming for the marking of these applications would be straightforward because the material to be typed does not require any use of specialist function or control keys; it is therefore easy for the software to check what has been typed against what should have been typed.
- 8.20 It is more difficult to assess at the computer the more advanced typewriting skills which require candidates to produce complicated layouts, to interpret handwritten corrections and to expand written abbreviations. Programming for marking these would be more difficult than for basic keyboarding, and would also need to take account of the accumulative effect of small errors. For example, if a word is omitted at an early stage in the passage, the software would need to allow for this in matching the subsequent typing against the perfect version.

- 8.21 One approach to computer assisted assessment of more advanced typewriting skills would be to provide a bank of paper-based test material, together with a floppy disc which would mark the output. The candidate would be required to type the material in any available word processing package, and to output the result on to an ASCII file. Marking could be for accuracy of keyboarding and correct interpretation of the source material. Some simple aspects of layout could also be assessed by this method.
- 8.22 In the future an alternative marking strategy might be to use a document scanner to read the candidate's hard copy output and to compare it with a 'perfect' version held on disc.
- 8.23 Assessment of word processing by computer delivered tests, and also assessment of the use of spreadsheets, databases, computer assisted design and computerised assessment, poses the problem that each software package uses different commands and conventions. Since the candidate will naturally wish to use the package to which he or she is accustomed, a separate assessment package would be needed for each type of software. This limitation has been accepted by the Aequitas series of tests which includes assessment packages for some of the more popular office software as well as the use of DOS. One particular advantage of computerised assessment of software use is that it is possible to monitor the candidate's keystrokes so as to check on the use of commands and functions, as well as to check the layout.
- 8.24 Employment Department action to develop or adapt Aequitas type tests is recommended. If these assessment packages proved useful and popular, the principle might be taken up independently, either by CBT producers or by the producers of the individual word processing database or spreadsheet packages. Further development might extend the use of assessment of this type to computerised accounting and design.
- 8.25 An alternative strategy would be to follow up a suggestion by Systems Applied Technology (SAT) for a generic package for teaching and assessing the use of database, spreadsheet and word processing packages. SAT's proposal is for a package which would teach the underlying concepts (for example, of 'fields' and 'global commands'), as well as their practical application. Because the system would not be specific to individual brands of software, only one version would be needed. However, it would only be fully appropriate for testing candidates who had learned using the same system; other candidates would wish to be tested on the software packages to which they were accustomed. As SAT is understood to have prepared a proposal, no separate proposal is included in this report.
- 8.26 It would almost certainly be possible to develop assessment software to assess and mark simple computer programming tasks. It is unlikely to be feasible, without great expense, to produce computer marking of complex programming tasks because of the many different solutions which exist to a complex problem. Computerised assessment of programming would therefore test basic understanding and the ability to solve problems, rather than high level programming skills. It would be necessary to define a fairly simple programming task for which there were only a small number of possible solutions. Marking would take account of correctness of syntax, appropriateness of screen layout and the correctness of the output for a given input.

- 8.27 Core skills in numeracy could readily be assessed by computer delivered tests. However, it is suggested that within the context of occupational NVQs, number skills would not normally be assessed in isolation but in conjunction with underpinning knowledge. They would therefore more naturally form part of knowledge tests. Even in the proposed new General NVQs it is likely that core skills would be assessed in context. No follow-up proposal is therefore given for development of computer delivered assessment in numeracy. It should be noted, however, that such tests would have considerable applications in the selection of students for courses or training programmes, and also for diagnosing weaknesses at the beginning of a course. Of the centres surveyed, 71% indicated that they would welcome tests designed to assess students/trainees on entry to the course.
- 8.28 Similarly, core skills in communication would not normally be assessed in isolation in the assessment of occupational NVQs. However, there are a number of occupations, particularly in the field of Business Administration, where the application of communication skills is important. There is potential for the development of computer delivered assessment in comprehension of both oral and written communication, and in the communication-related skills of spelling and identification of correct usage of English. There are also technical occupations in which comprehension of technical material is important.
- 8.29 It would therefore be beneficial to develop material for the assessment of candidates' communication-related skills at the computer. This would probably require a split screen so that the stimulus text could be displayed and scrolled independently of the related questions. For oral comprehension an audio facility would be required. One further development which might be attempted is to assess the skills involved in taking telephone messages by asking candidates to listen to a recording of a message and to select those phrases which are both correct interpretations of part of the message and relevant items to include when transmitting the message to another person.
- 8.30 Similar techniques could be used to assess the comprehension of foreign languages in a business or leisure services context.
- 8.31 Some more specialised aspects of communication skills could potentially be assessed, at least in part, by computer. Examples are:
- map reading, which features in NVQs for army personnel and for electricity board linesmen
  - comprehension of engineering drawing, which is relevant to a wide range of engineering personnel
  - morse code transmission and reception, needed by electronic communications operators.

- 8.32 In occupations where information skills are important, computer assessment of these skills might be possible. The simplest method would be to provide a computer delivered objective test related to commonly used reference books, to which the candidate would be allowed free access. It would, of course, be essential to ensure that both assessment and candidates used the same edition of the reference works. A more ambitious development would require the candidate to access reference material held on CD-ROM; the software could track the stages by which the candidate accessed the information.
- 8.33 Problem solving, is a technically feasible area for computer delivered assessment. It is recommended, however, that, within occupational NVQs it should not be assessed in isolation but in relation to relevant occupational tasks. One practical application of problem-solving skills is in simulations of fault-finding, and these are discussed in Section 10; other types of practical problem solving could probably also be assessed by simulation techniques.
- 8.34 Techniques already used in computer based training could be used to develop computer delivered assessment of skills related to accounts work, such as the completion of petty cash books and ledgers to the standard required for Business Administration Levels 1 and 2. Candidates would be tested on their understanding of the various recording formats used, the entry of figures into the correct columns, correct totalling and so on. With experience at NVQ Levels 1 and 2 it would probably be possible to develop computer delivered tests of accounting skills at higher levels.
- 8.35 Some CBT packages and computer delivered selection tests already assess skills related to filing; particularly the ability to put items into alphabetical order or into other categories. Such exercises might form part of a larger assessment package for Business Administration, together with tasks related to sorting mail and 'in tray' exercises.

#### Possible markets

- 8.36 As described in Section 7, the potential market for computer delivered skills tests includes awarding bodies, centres and TECs.
- 8.37 The most obvious market for many of these applications is in the general area of Business Administration, where candidate numbers are large, and in Information Technology. However, a number of individual skills would also be relevant to areas such as Retailing and Hotel and Leisure Work.

#### Possible follow-up action

- 8.38 It is recommended that one aspect of all follow-up projects should be that the software produced for computer delivery of skills assessments should be capable of handling both fixed assessment tasks and also material produced by the task generation techniques described in Section 6 of this report.

8.39 Outline project specifications are provided for the following follow-up projects:

- development of computer delivered tests of basic keyboarding (G8.1)
- further development of G8.1 to provide a computer delivered test of audio typing
- investigation of computer marking of more advanced typewriting skills (G8.3)
- development, or adaptation for the UK market, of computer delivered tests of word processing, database and spreadsheet use (G8.4)
- exploration of computer assessment of programming skills (G8.5)
- development of computer delivered assessments of communication skills (G8.6) and of accounts-related skills (G8.7).



## SECTION 9: DELIVERY OF KNOWLEDGE ASSESSMENTS

### Main features

- 9.1 Computer delivered knowledge assessment offers the possibility of testing at the computer, candidate's underpinning knowledge and understanding and their ability to apply this knowledge and understanding to given situations. If maximum benefit is to be gained, the questions should be posed in such a way that the answers can be marked objectively by the computer. However, this does not mean that the technique is limited to multiple-choice questions. A variety of other objective formats are available, including multiple-response (where more than one of the listed options is correct), matching, insertion of words in text or deletion of unnecessary words and direct input of numerical answers.
- 9.2 Computers could also provide a mechanism for storing and delivering assessments which require marking by examiners, with the candidate responding either at the computer or to a printed question paper. Because this offers fewer benefits in terms of gains in efficiency it is given less consideration in this report.
- 9.3 Computer delivered knowledge assessments can make use of 'stimulus material' in the form of paragraphs of text, computer graphics, audio input and video.

### Contribution to improving access, efficiency and effectiveness

- 9.4 Contributions to the improvement of access, efficiency and effectiveness for all computer delivered assessment are described in paragraphs 7.7-7.17 of Section 7. The particular benefits of testing knowledge by computer delivered means are that this provides a way of extending the scope and range of the total assessment process without the need for the candidate to demonstrate performance in a large number of different situations. Computers may also assess knowledge and understanding which cannot be inferred from performance on the job.
- 9.5 It is not suggested that assessment of underpinning knowledge, whether by computer or otherwise, should be seen as an alternative to workplace demonstration of competence. The value of knowledge assessment is in extending the range of the assessment.

### Current availability and use

- 9.6 There are a small number of existing computer delivered tests of knowledge in specific subject areas. The only such tests identified which actually lead to a qualification are the two City and Guilds schemes *Inside Information* and *Inside Science*. Each of these is linked to a BBC series and book, and a candidate who completes the associated multiple-choice test on a computer qualifies for a certificate.

- 9.7 There are a number of existing tests with specific content which do not lead to a certificate. City and Guilds has a knowledge test on the Accreditation of Prior Learning process which forms part of a resource pack on APL for verifiers. The Hairdressing Training Board has knowledge questions relevant to its NVQs on a computer-based system which centres may obtain, but it is understood that these are used for formative purposes rather than as evidence towards certification. GES developed a suite of knowledge tests relating to basic electronics to provide assessment of trainees on entry to Skillcentre courses for the then Skills Training Agency; it is understood that these are not now widely used. A battery of adaptive tests for college entrants is produced by Educational Testing Services for use in US community colleges.
- 9.8 There are several shell software systems which may be used to construct and deliver multiple-choice and other tests in subjects determined by the user. The most popular and inexpensive system known to be available in the UK is Question Mark, which has a wide range of users in colleges and amongst training providers. Question Mark has an option which enables graphics to be incorporated in the questions. A new option (Talking Question Mark) allows audio input and may be used as part of the testing of language comprehension or of the significance of particular sounds (e.g. heart murmurs, engine noises).
- 9.9 Other shell systems on which information has been obtained are Gapmaster (in which the candidate has to select the correct word to fill a gap in a sentence), the Assessment Toolbox associated with the Standards candidate recording software, Examiner (developed in the United States for the Stock Exchange but available on sale in the UK) and Microcat (not, as far as is known, on sale in the UK). Several firms produce selection or general ability tests in computer delivered format, and these could be adapted for delivery of knowledge tests, but with the exception of Selby-MillSmith these firms do not appear to be interested in making their software available as a shell.

#### On-going research

- 9.10 Astra Training Services (formerly the Skills Training Agency) is piloting software developed by a European consortium for analysis of candidates' responses to multiple-choice questions. This is to be extended in future stages of the project to provide computerised banking and test delivery. Astra is particularly concerned with the piloting of carpentry and joinery questions, but the software is produced as a shell, and other European partners in the consortium are working on other subject areas.
- 9.11 System Applied Technology is developing software for computer delivery of knowledge tests for use by the British Polymer Training Association in NVQs for Plastics Processing.
- 9.12 It is understood that the Royal Institute of Public Health and Hygiene is holding preliminary discussions on the possible authorisation of a computer delivered testing system for one of its examinations.
- 9.13 The Associated Examining Board is also understood to be planning to introduce computer-delivered testing, for a vocational rather than a GCSE examination.

- 9.14 Research on adaptive testing presupposes the availability of a computer to enable the questions delivered to be tailored to the individual candidate. ETS and Sellby-MillSmith software allow adaptive delivery.

#### Potential for further development

- 9.15 Although there are a number of systems in existence, it appears that there is no generally available UK-produced shell software which would fully meet the needs of an awarding body wishing to use computer delivered assessment. 'The Examiner', produced in the USA but marketed in the UK, appears to fulfil most of the requirements but would possibly need adaptation for situations in which the tests were used outside the organisation which produced them. Question Mark has many advantages, not least its low price, but it is understood that it does not use item analysis data in the selection of questions for presentation. Other existing systems are mostly specific to one or two subject areas or do not include sufficient graphics capabilities.
- 9.16 It is highly desirable that if computer delivered assessment is to be used on a large scale within National Vocational Qualifications, the software used should be, as far as possible, similar across all subject areas and awarding bodies. Candidates may well take assessments from more than one awarding body in the course of building up the required units for a qualification, or adding to their qualifications after experience; this is a feature of the flexibility of the NCVQ system which NCVQ and the Employment Department wish to promote. It would be confusing for candidates if software required different procedures, for example for recording responses to questions. There is therefore a strong case for the Employment Department to fund the production or adaptation of shell software which would meet the needs of both large and smaller awarding bodies and (perhaps in a cut-down version) individual centres. Producers of existing software could be invited to tender for this work, which would then build on their existing expertise.
- 9.17 Development of such a system should be preceded by consultation with interested parties to refine proposals for minimum requirements for such a system.
- 9.18 Work on the software should proceed in parallel with work on the administrative and security procedures which are necessary to ensure that the assessment system will work efficiently and in a manner which will reassure users as to its fairness and reliability. These aspects, together with the infrastructure which is needed, are partially common to all computer delivered assessment and are discussed in Section 7.
- 9.19 Experience has not yet been gained in large-scale computer delivered assessment of knowledge, and it would be desirable to trial computer delivered knowledge tests in a subject area where there are substantial numbers of candidates. This might be in an area such as engineering, where banks of questions already exist, although the suitability of the questions would need to be reviewed in the light of the development of new NVQs. Alternatively, initial development might focus on one or more NVQs where there is little previous experience of assessment and where the assessment system needs to be developed from scratch; this would have the advantage that it would be less influenced by pre-conceived ideas, but the disadvantage would be that there would be less existing work on which to build.

- 9.20 Provision for a suitable software shell will not by itself guarantee the quality of the assessment material inserted into that shell. Indeed, computer delivery may give a false impression of quality to an inherently poor test. Experience of examination design in general, and of multiple-choice questions in particular, shows that for inexperienced question writers there are a great many pitfalls in the task of producing suitable assessment material. City and Guilds, when introducing multiple-choice questions in the late 1960s, embarked on an expensive programme of training of 'item writers'. While City and Guilds' existing expertise should be harnessed in any new developments, there will also be a need to gain experience of setting the different question types which will become possible with the new computer delivery software. In any new NVQ for which computer delivered knowledge assessments are introduced, it will be necessary to provide training for the question writers; smaller awarding bodies in particular will need help in this area.
- 9.21 Even after a shell had been developed for use by awarding bodies (which would itself take some time), Question Mark might continue to fulfil a need at centre level because of its low cost. For many NVQs there may be no centrally produced knowledge assessment material for some years to come, or indeed ever, and centres are likely to continue to use Question Mark for entry level and formative assessment. It is therefore suggested that the Employment Department might consider holding workshops on assessment quality for users of Question Mark (and perhaps of any other commonly used packages). Such workshops could also provide extremely valuable feedback as to the types of situation in which computer delivered knowledge assessments are being used and the problems which are being experienced.
- 9.22 There is scope for much more development in relation to the use of audio and video stimulus material in connection with knowledge testing. It is notable that the majority of the interactive video training packages which have been viewed in the course of this project make very little use of video in their test sections; most of the tests viewed in IV packages could be delivered by straight-forward CBT. One example of a selection test making full use of linear video for testing observational skills has been seen. It seems likely that there will be a number of vocational areas where satisfactory computer delivery of knowledge assessment would require the use of, at the very least, high quality computer graphics, and possibly video material. Moving pictures could be particularly useful in occupations which relate to the care of animals or people. Both linear and interactive video should be considered as possible stimulus material.
- 9.23 Another area which merits investigation is the use of audio to supplement the text material in the assessment of special needs candidates, specifically the partially sighted and those with literacy problems. There are some occupations in which the level of literacy directly required by the work is very low, and candidates who have the necessary subject knowledge would be disadvantaged if presented with computer delivered questions demanding a level of reading skills which they do not have. Existing Level 1 NVQs get round this problem by using oral questioning. An alternative would be to record the text of the questions and play it to the candidate when the text was displayed. This technique is used by the ETS Procare software. The use of audio with text is also being developed for use in the teaching of adult literacy students by Cambridge Training and Development.

- 9.24 Possible developments in the use of item analysis to provide feedback for quality control purposes are discussed in Section 14.

#### **Adaptive testing**

- 9.25 Considerable interest has been expressed in the use which might be made of adaptive testing in assessment for NVQs. The use of computer delivery of assessments opens up a possibility which scarcely existed in the days of paper based tests for tailoring the questions presented to the candidate according to the responses which he or she has already made. Research on adaptive testing to date has focused particularly on its use in tests of general abilities or broad areas of attainment.
- 9.26 One main use of adaptive testing is in varying the difficulty of the questions presented to the candidate, so as to provide a more exact measure of his or her ability. A conventional fixed content test does not provide accurate measurement of candidates' ability at the extremes of the range since it is primarily geared to the needs of candidates of average ability. By varying the questions presented to those candidates who showed themselves to be not of medium ability, the test can provide a more accurate measurement of their place on the scale. Adaptive testing of this sort is more relevant to selection tests, course entry tests and formative testing than to testing in relation to a pre-determined standard. Although testing candidates on entry to the course was an area in which the centres surveyed expressed a considerable interest, it is not directly relevant to the issue of testing for accreditation, and has therefore not been pursued further in this project.
- 9.27 An alternative use of adaptive delivery is in mastery testing, where the aim is to shorten the test by presenting each candidate with questions which will provide the maximum information as to his or her mastery or non-mastery of the relevant body of knowledge. There are two ways by which this may be achieved, Adaptive Mastery Testing and Sequential Mastery Testing. The basic underlying theory of each involves the probability of any given score accurately reflecting the candidate's mastery or non-mastery. When the candidate's responses to the questions presented so far result in a score which puts him or her in either of these two categories with a sufficiently high degree of probability, then the test may be terminated (to use a very simple example, if the 'cut score' or pass mark is 35 out of 60, and the candidate answers the first 35 questions correctly, then there is little point in administering the rest of the questions). The main advantage of Adaptive Mastery Testing over non-adaptive testing is the potential saving in candidate time, and hence also in the amount of time the computer is in use.

- 9.28 There are two potential disadvantages of Adaptive Mastery Testing for use in NVQs. One is that the sample size required for pretesting is normally larger than for traditional item analysis, typically several hundred candidates. However, in a number of important vocational areas this would not present any problem. Another possible problem is that adaptive testing theory assumes that the test is 'unidimensional', i.e. that there is only one ability or trait being measured. This is not necessarily the case with vocational qualifications where it has traditionally been assumed that it is necessary to have a representative syllabus coverage in each test. For example, if the content of the qualification includes resistors, capacitors and transistors, then the candidate should be tested on his knowledge of all three, and it is not assumed that because he knows about resistors he also necessarily knows about transistors. An adaptive test which did not cover all topics could well lack credibility.
- 9.29 There is potential for further investigation of the possibilities of adaptive testing, but necessary prerequisites are an administrative infrastructure and system for the delivery of computer based tests, identification of suitable occupational areas where the technique could be trialled, and an infrastructure for pre-testing the questions. The most appropriate subject area in which to work would be one in which some elements of underpinning knowledge occur in two or more levels of NVQs, with each level becoming more demanding in the degree of knowledge and understanding required. It would then be possible to set up an adaptive test which could accredit students at one of two or more levels according to their responses. A possible subject area might be the understanding of logic gates in electronics.
- 9.30 Because this work would pre-suppose other work on infrastructure and delivery mechanisms, as well as an appropriate NVQ area in which to work, it is suggested that further development of this aspect of adaptive testing should be seen as falling within a second stage of development of computer assisted assessment.
- 9.31 Although the need to assess underpinning knowledge is clearly set out by NCVQ, there is still considerable misunderstanding about the potential of knowledge tests in general and objective tests in particular in assessment for NVQs. This needs to be overcome before the greater subtlety of adaptive testing can be introduced.
- 9.32 It has been suggested that a form of adaptive testing (although not falling within the stricter definitions used by psychologists) might be employed to extend the range of assessments delivered to an individual candidate. The assessor would observe the candidate demonstrating competence in one particular situation and would then, in some way, interact with the computer to ensure that the knowledge questions delivered to the candidate covered other aspects of the range. Although this would reduce the time which the candidate needed to spend on the knowledge assessment, since he or she would not be asked questions on areas covered in the workplace observation, the savings would probably be small in relation to the cost and effort of setting up and implementing the system. It is suggested that work on these lines could more usefully be undertaken after further work has been done on guidance for assessors and the evaluation of combinations of evidence, as discussed in Section 11.

### Possible markets

- 9.33 Markets for computer delivered assessment would be either centres and TECs, or awarding bodies and Lead Bodies, as discussed in Section 7. Knowledge assessments can apply across all subject areas; use of video may be found most appropriate in occupations such as agriculture, horticulture and caring.

### Possible follow-up action

- 9.34 A number of follow-up projects are proposed in the light of the discussion earlier in this section, and outline project specifications are provided for the following:
- a shell delivery system for computer delivered knowledge tests (G6.1)
  - discussions on the necessary infrastructure and security procedures for computer delivery of assessment (G7.4)
  - development of a bank of knowledge assessment material for computer delivery for one subject area, and trialling together with the associated administrative and security procedures; results to be used for accreditation where appropriate (G9.1)
  - similar development of a bank of material which could be relevant to more than one NVQ in related areas, for example welding for both engineering and shipbuilding (G9.2)
  - exploration of the potential for setting questions related to existing video material (G9.3)
  - joint development of the Learning Technologies Unit of an interactive video package for both training and assessment (G9.4)
  - exploration, in connection with special needs advisers within TEED, of the potential for using audio in conjunction with computer delivered assessment for special needs trainees (G9.5)
  - hosting of one or more use days for Question Mark users, to identify problem areas and to highlight the need for quality (G9.6)
  - training sessions for awarding bodies and Lead Bodies planning to introduce computer delivered knowledge tests (G9.7).

## SECTION 10: COMPUTER SIMULATION

### Main features

- 10.1 Computer simulation involves an interaction between the user and the computer in a way which mimics real life. At each stage the computer's response depends upon the inputs made by the user. Examples of the use of computer simulations in training include operation of plant and equipment, fault-finding exercises, and crisis management.
- 10.2 Computer programs which mimic the behaviour of commercial software (e.g. spreadsheets) for teaching or assessment purposes have some features in common with simulations, but are less adaptive. They have been covered under the heading of skills assessments (Section 8).
- 10.3 Large-scale industrial simulators, which use computers in conjunction with other hardware, are not considered in detail in this report. They are specific to individual industries and are likely to be used for teaching and testing only where the economics of the situation or the critical importance of worker competence make it worthwhile for the industry to finance their development.
- 10.4 Some recent developments in computer simulation have emphasised the use of interactive video, but use of video is not essential for simulations; good computer graphics may be sufficient.

### Contribution to improving access, efficiency and effectiveness

- 10.5 An important benefit of using computer simulations in assessment is that they can extend the range of the assessment by dealing with situations which occur infrequently in practice or are dangerous or expensive to set up for assessment purposes. Crisis management is one example where simulations are preferable to real emergencies.
- 10.6 Use of computer simulations can also shorten the overall timescale of an assessment and hence improve its coverage, because the computer can provide quicker feedback than would be obtained in a real situation. Examples are plant and equipment operation and fault-finding, where the computer can report the results of 'tests' or altered machine settings, requiring the candidate to respond to the new situation.
- 10.7 Computer simulations can thus improve the **effectiveness** of assessment by providing a wider coverage than would be available from workplace observation alone, and a more realistic assessment than would be obtained from knowledge testing.
- 10.8 In the centre survey, 54% of respondents said they had 'considerable' or 'minor' difficulty in arranging candidate assessments which covered dangerous or emergency situations, and 63% reported difficulty covering situations which occur rarely in practice. For industrial and Lead Bodies the figures were 69% and 82% respectively.



## Current availability and use

- 10.9 There is some existing use of simulation techniques in computer based training, but none of the applications are used to assess candidates for accreditation, and in many of the software packages viewed the section on assessment is relatively weak.
- 10.10 Two particularly successful CBT packages on fault-finding were viewed during the project. One, developed by Rediffusion and the Road Transport Industry Training Board with some TEED funding, was for teaching and testing fault-finding skills in connection with electronic ignition systems. In the test section, the user is given the customer's report on the fault symptoms and then proceeds to check or test possibly relevant points on the vehicle as a whole (e.g. 'petrol in tank') or in the detail of the fuel injection or electronic ignition system. The result of each check is reported on screen, and the times required for each are added to a running total. At the end, the user's time to solve the fault is reported, together with the time of an 'expert'; thus a user who makes unnecessary or irrelevant checks incurs a time penalty.
- 10.11 The other successful fault-finding package viewed was an interactive program produced by Cleveland Open Learning Unit. This uses the context of fault-finding in an electronic circuit, although the necessary electronic knowledge is provided and the main aim is to test the candidate's ability to use logical fault-finding techniques. The screen displays a diagram of a branching electronic circuit with information about the observed faulty output, and the candidate is asked to specify the points at which tests are to be made in order to determine which component is faulty. After each 'test' the results are reported and the candidate then selects another test point, until he or she is able to specify where the fault is. The software logs the tests made, and reports on whether any unnecessary tests have been made and whether the most efficient testing sequence was used. By contrast, another electronic fault-finding package viewed made little use of the potential interactivity of the system.
- 10.12 A number of computer-based training packages exist to train operators in the use of specific types of industrial plant or equipment; one active producer is Rediffusion. Examples are packages to teach plastic injection moulding machine setting, techniques for engineers controlling the national grid, and control of the processes for sterilising cans in a food processing plant. Several firms produce CBT to provide the first stage of training for aircraft pilots converting to a new type of aircraft; use of CBT reduces the time which they need to spend in full-scale simulators.
- 10.13 Similar techniques are used in a package to help boiler plant controllers to analyse boiler water and assess resulting information and another to help chemical technicians to assess the results of tests using techniques such as spectroscopy.
- 10.14 The Employment Department have funded, or part-funded, two training packages concerned with crisis management. ICCARUS (control of large fires) and a package on crowd control for police officers.

- 10.15 A series of CBT packages which are not simulations in their present form but suggest possible applications, are produced by a company called Clearsoft, based at a former Skillcentre in County Durham. Each of these packages provides a diagram of an industrial or office situation and requires the candidate to spot the hazards by moving a cursor to the position of the hazard and then inputting one or two words to explain the nature of the hazard. The main use of these packages at the present time is as an aid to teaching and a stimulus to class discussion.
- 10.16 Use of large-scale simulators exists in air pilot training, and National Power has a simulated grid control room at its training centre in Leeds.

#### **On-going research**

- 10.17 One area where work is underway to use simulation techniques for purposes of accreditation is a project being undertaken by System Applied Technology for the British Polymer Training Association, with partial funding from the Employment Department. In this system, the screen will display a schematic diagram of an injection moulding plant and describe the fault(s) on the components being produced. The candidate is required to identify the checks to be made to discover the cause of the problem. The marking system is not yet fully determined but will give credit for use of a logical testing sequence and the inclusion of essential safety checks.
- 10.18 Work is also being undertaken for Bus and Coach Training Ltd by Rediffusion, again with Employment Department funding, to assess the reaction of bus and coach drivers to crisis situations such as a road accident. Development is at an early stage, but it is understood that the candidates will be given as few directions as possible on what to do, and required to use a mouse to indicate on the screen areas of potential help or in connection with which action might be taken (e.g. clicking on a telephone box to indicate the intention to summon help).
- 10.19 PRD is developing a package to teach and assess the management of crises in large PVC plants.
- 10.20 The Employment Department is currently commissioning a package relating to hazard spotting in building sites.

#### **Potential for further development**

- 10.21 Many of the techniques already used in computer based training, have potential applications and possible further developments in computer assisted assessment. However, because the technology is likely to be relatively expensive, relying on interactive video or high quality graphics, it is necessary to examine potential applications to ensure that the benefit to be gained will justify the costs of development.

- 10.22 Use of computer simulations to assess fault-finding skills is possibly the area which has the greatest potential for short-term development. Fault-finding skills are important in a wide range of occupations (including electronics, plumbing and heating and motor vehicle maintenance), and computer assessment offers the possibility of testing candidates on a wider range of situations than would necessarily occur naturally in the workplace over a relatively limited period. The potential interactivity of a computer system means that the assessment can be more realistic and more responsive to the candidate's input than a paper and pencil test would be.
- 10.23 It is understood that computer programs have already been developed in the United States to simulate the process of medical diagnosis. It might be feasible to use these techniques at NVQ Levels 2 and 3 to test candidates' response to potential illness in caring and animal care occupations. While candidates at this level would not be expected to make professional diagnoses, they need to be able to recognise the signs of conditions which can be corrected without professional help and those which require them to summon help.
- 10.24 Computer simulations of the operation of processing plant or equipment could be adapted for use in assessment, as could programs simulating the conduct and interpretation of chemical analysis.
- 10.25 Simulations involving crisis management also have potential for use in assessment purposes. However, if the simulations are to be realistic they are likely to require consideration of a number of factors and in practice decisions will be influenced by many related considerations. Programs simulating conduct of a crisis would probably therefore best be used as a source of discussion material to form the basis of an oral questioning session with a human assessor. Any attempt to reduce each stage to a list of possible options from which the candidate had to choose only one, without any explanation of his or her choice, might lead to a trivial test. Considerable work would be required to define all possible answers in order to allow computer marking of free text input.
- 10.26 There may be other problem-solving situations which could be computerised to form the basis of assessment. For example, the project team has seen, in another context, a paper-based 'maze exercise' for secretaries, in which the user reads a description of the situation, decides which action to take first, and then turns to the appropriate page of the booklet for feedback and the next stage of the problem. This has obvious potential for computerisation, and could perhaps form the basis of assessment of some of the secretarial competences at NVQ Level 3.

- 10.27 There are possible applications of the hazard spotting idea in computer simulation for purposes of assessment. Many occupational standards require candidates to be able to identify safety or hygiene hazards in the workplace and take appropriate action, but these are difficult to assess because the appropriate hazards may not occur in the workplace. Knowledge questions tend to direct the candidate too obviously to the correct answer, except where they relate to an understanding of the reasons for correct/safe procedures. What is really important is to try to assess the candidate's ability to notice and take action in relation to safety hazards. It is therefore suggested that an area for investigation should be the feasibility of producing an interactive video in which the candidate was asked to watch a film and to record any instances where a safety hazard was spotted. This would be appropriate, for example, in relation to workshop and site safety and in residential care situations. Similar techniques might be used to test security officers' observation of apparent breaches of security. Initial testing of the material could be by use of linear video.
- 10.28 More attention needs to be devoted to the measurement aspects of these new assessment techniques, including the processes and outcomes to be assessed, scoring methods and techniques for analysing candidate responses.

#### Possible markets

- 10.29 Possible markets for computer delivered assessment via simulation would be mainly in subject areas where the interactive feature is particularly desirable, but this would cover a fairly wide range of technical and managerial occupations.
- 10.30 Markets might be either centres or awarding and Lead Bodies, but because many of the applications will be specific to a relatively small occupational area it is likely that the Lead Bodies will wish to play a larger part in the development of such applications than might perhaps be the case with skills or knowledge assessment.
- 10.31 Where computer delivered assessment by simulation needs to include interactive video or even simply high quality graphics, the higher development costs will be accompanied by higher costs for the users in providing the necessary equipment. Computer simulations will therefore not be cost-effective over such a wide range of applications as computer delivered knowledge assessment, but might be chosen as the only feasible assessment method for some elements.
- 10.32 Because of the cost implications, it would be desirable to explore the potential for linking computer assessment packages with training packages. This might be done in one of three ways:
- 'licensing' existing software to indicate that a candidate who had successfully worked through a package could be regarded as having provided the necessary evidence of competence in relation to one element or part of an element
  - developing assessment software which used graphics or video material already developed for a training application
  - joint development of learning and assessment material.

- 10.33 Linking assessment with learning material would be governed by the 'Acceptability Criteria' proposed in Section 7 (see paragraph 7.39 and , project proposal G7.6). Experience in developing one or two joint packages is likely to lead to a refinement of the criteria.
- 10.34 Any projects to link assessment and learning material should also be alert to the possible danger of distorting the learning material, as discussed in paragraph 7.24.
- 10.35 Simulation packages for assessment should be developed so as to allow a large number of different assessment problems to be generated from a restricted number of graphic images or video sequences.

**Possible follow-up action**

- 10.36 Outline specifications for follow-up projects are provided for:
- Development of an assessment package for fault-finding in perhaps two specific subject areas, one for large-scale plants such as chemical plants, the other for smaller equipment such as heating or motor vehicle maintenance. (G10.1)
  - Drawing out of lessons to be learned and identification of those aspects of the software which could provide the basis of a 'shell' for fault-finding assessments. (G10.2)
  - Discussion with lead and awarding body experts in two or three occupational areas to consider the feasibility of 'licensing' existing software for assessment purposes. (G10.3)
  - Joint development with LTU of a pack for both training and assessment in an area of plant or equipment operation, perhaps pharmaceutical processing or glass production). (G10.4)
  - Project to develop an assessment package on the spotting of hazards. (G10.5)

## SECTION 11: AIDS TO DECISION-MAKING

### Main features

- 11.1 There are three areas in which help may be needed for assessors and verifiers involved in making decisions about the competence of individual candidates and in which computer assistance could play a part.
- 11.2 Help is needed for assessors in deciding whether individual demonstrations of competence are adequate, for example whether they are at the correct standard and based on tasks of an appropriate degree of complexity. Although the statement of elements and performance criteria should in theory provide sufficient guidance to those undertaking assessment, it is inevitable that there will be points of uncertainty until the standards are well established and there has been experience of their implementation.
- 11.3 Guidance may also be needed by assessors and verifiers who have to combine evidence of candidate competence from a variety of different sources which may include both workplace observation and questioning of underpinning knowledge. The aspect which is most likely to cause problems is where the range statement provides for competence to be applied in a number of different circumstances.
- 11.4 Assessors may also require help in structuring interviews and conducting oral questioning of candidates in such a way as to elicit all the necessary evidence of competence and underpinning knowledge.

### Contribution to improving access, efficiency and effectiveness

- 11.5 Computerised assistance to assessors and verifiers in these aspects of decision-making is likely to improve the **quality** and **consistency** of the assessment by helping to ensure that all are working to common standards. Coordinating the standards of assessors and monitoring the performance of centres were regarded as posing 'considerable' or 'minor' problems by almost 80% of industry bodies. A significant minority of centres felt that there was considerable scope for improvement in the information and guidance provided in relation to different aspects of assessment. The percentage of centres experiencing 'considerable difficulty' in deciding whether individual candidates had fulfilled the requirements for a particular unit or element was 28%, with 43% experiencing minor difficulty; 11% were experiencing considerable difficulty (44% minor difficulty) in weighing the value of different types and combinations of evidence.
- 11.6 The major contribution of aids to decision-making would be in improving the quality of assessment, but there would also be some improvements in **efficiency** since assessors would waste less time in uncertainty over how to rate candidates, and might also be able to shorten the interview process. There would be further savings if similar guidance were made available to candidates as discussed in Section 5 (para 5.45) so that the candidate played a larger part in the selection of suitable evidence.

#### Current availability and use

- 11.7 Guidance for assessors and verifiers is provided in paper-based forms in conjunction with individual sets of standards and computerised guidance systems and aids to decision-making exist in other spheres, but there has been relatively little experience of combining the two into computerised guidance for assessors.
- 11.8 Some work has been done in connection with Accreditation of Prior Learning, as discussed in the next section. The 'Capability' system is intended mainly for use by individual candidates, but the FEU expert system programs which have been developed on a trial basis are intended to provide help for the assessor attempting to evaluate the prior learning evidence of candidates for qualifications in electrical installation work and motor vehicle maintenance.
- 11.9 The RSA Examinations Board has published a helpful guide on APA (the Accreditation of Prior Achievement), which provides, in paper-based form, examples of acceptable evidence for a small selection of units of competence.
- 11.10 The Central Council for Education and Training in Social Work (CCETSW) has developed an interactive video to provide training for its workplace assessors in the assessment of candidates for caring occupations. A linear video version of the case studies is also available.
- 11.11 During the pilot year of the Certificate in Open Learning Delivery (1988-89), the assessor training sessions included a video of a candidate demonstrating relevant competences. This was not entirely successful because the quality of the video was rather poor.
- 11.12 NCVQ is about to begin production of interactive video material for use in training assessors to the standard of the Training and Development Lead Body units.
- 11.13 In the field of human resource management, work has been done to attempt to define the competences required for individual jobs and to match these against databases of the skills of the existing workforce. The most developed system is Peodesy (formerly known as PARYS); this is a large system with a considerable number of individual modules, including those required to match job requirements against applicants or existing staff.

#### On-going research

- 11.14 The Employment Department is funding a project being undertaken by Dr David Bartram to investigate issues relating to the sufficiency of evidence. This project which is still at a very early stage, should produce useful, general guidelines for those developing standards and associated assessment requirements.
- 11.15 The Institute of Manpower Studies at Sussex University has experience of advising clients on the application of computer based systems to human resource management problems, including the decision-making process and including systems which can incorporate weightings for different types of evidence.

## Potential for further development

- 11.16 It is likely that assessors in most NVQs would benefit from the provision of more guidance and specific examples of acceptable and insufficient candidate evidence, together with a 'bank' of assessment decisions made. The ability to access a resource bank of this type would improve the consistency of decision-making amongst all the assessors for any individual NVQ. A possible mechanism for storing and disseminating this information would be to compile a database for each NVQ, cross-referenced according to the individual units and elements, and providing information about types of evidence accepted or not accepted and assessment decisions made. Assessors encountering problems not already covered by the database would be expected to refer them to the verifier who would ensure that the agreed decision was entered on the database. Dissemination could be by provision of updated floppy discs at regular intervals. This mechanism would avoid the need for continual face-to-face meetings of assessors after initial training had been provided.
- 11.17 It might be necessary to support this database by the provision of paper-based or video examples of acceptable or insufficient evidence. For ease of dissemination this should be avoided wherever possible, but in occupational areas like caring it is likely to be beneficial to provide video or interactive video showing, for example, candidates interacting with clients.
- 11.18 There have been a number of suggestions that an expert system could help assessors and verifiers to combine evidence of candidate performance in order to reach a decision as to whether the candidate has demonstrated competence. However, staff at the Institute of Manpower Studies emphasised that an expert system is only as good as the information it contains, and that a necessary precondition would be a programme of work to improve and validate the level of consistency between assessors.
- 11.19 Another difficulty in adapting the sort of expert systems used for personnel decisions is that these take 'scores' for a number of factors, each of which are allocated weightings. By contrast, NVQ assessments would be a series of pass/fail judgements, all of which would merit the same rating.
- 11.20 There might, however, be scope for a simple decision support system which could help assessors and verifiers to make consistent decisions based on combinations of performance and knowledge evidence relating to different competence elements and aspects of the range. Situations which particularly lend themselves to this approach would be NVQs, where the range statements cover a number of items and individual candidates present different combinations of evidence, with one or two items relating to performance in the workplace and others consisting of knowledge evidence across the range. A combination of evidence sufficient for accreditation might require a minimum number of situations to be covered by workplace observation and other aspects of the range to be covered by knowledge assessment.



- 11.21 The sufficiency of evidence project currently being undertaken could be followed by one or two subject specific projects. These would be intended to produce guidance on acceptable combinations of evidence in a form which could be entered into a rule-based decision support system. The discipline of defining the evidence rules with sufficient precision might lead the subject experts to specify the requirements of the qualification much more clearly than is done at present. This could provide useful insights for other occupational areas, even where there was no intention to put results into a computerised system.
- 11.22 If this approach proved successful, it could be adapted into a computerised interview guidance system. This would help assessors to ensure that the oral questioning they conducted covered sufficient items in the range and statements of underpinning knowledge to supplement the evidence obtained from workplace observation.

#### **Possible markets**

- 11.23 The market for any aids to decision-making would be the awarding body responsible for each award who would need to provide these aids to all the assessors and verifiers concerned with the award. Awarding bodies would need either to be assisted with initial development or to be convinced that such aids were cost-effective.
- 11.24 Experience of developing computerised aids to decision-making in one or two NVQs should make it possible to identify the common factors applicable to all occupational areas and hence to lead to the development of a shell software system which could be customised for individual NVQs. This would have the effect of reducing the development costs required for individual NVQs.
- 11.25 Initial development should, however, investigate the usefulness of computerised systems for assessors. It might be that paper-based versions would fulfil the requirements and would be more convenient in use.

#### **Possible follow-up action**

- 11.26 Follow-up projects are proposed to explore these ideas in the context of individual subject areas. Outline project specifications are provided for:
- a database of assessment decisions and examples of acceptable/unacceptable evidence, as a source of help for assessors (G 11.1)
  - development of a decision support system to help assessors judge different combinations of evidence (G 11.2)
- 11.27 If the decision support system proved successful, it could be further developed into an interview guidance system at a later date.

## SECTION 12: ACCREDITATION OF PRIOR LEARNING

### Main features

- 12.1 This section discusses computerised guidance systems which can help the candidate considering applying for a qualification via the Accreditation of Prior Learning (APL) process to identify his or her relevant past experience. The computer system helps the candidate to relate experience to the competence elements within the qualification and to identify sources of supporting evidence. When the candidate has obtained the appropriate evidence the portfolio is then discussed with an assessor or counsellor.
- 12.2 Similar systems encourage the candidate to consider his or her own strengths and weaknesses and to give a self-rating for each competence element which is then compared with the ratings of the candidate's supervisor or line manager.
- 12.3 Other computerised guidance systems for use in APL are intended as an aid to the assessor, rather than to the candidate, and provide a guide to the points to be considered in evaluating the candidate's level of prior knowledge.

### Contribution to improving access, efficiency and effectiveness

- 12.4 As discussed in Section 11, systems intended to guide the assessor are mainly advantageous in improving the quality and consistency of the assessment process. By contrast, systems intended to provide help for the candidate can greatly reduce the time which the mentor or assessor needs to spend with the candidate, thus **reducing costs and improving efficiency**. Beneficial side effects are that the APL process becomes more affordable and less daunting to the individual candidate, so **improving access**, and that the candidate assumes responsibility for the process.

### Current availability and use

- 12.5 The 'Capability' program developed by Crosskeys and Newport Colleges allows APL candidates for motor vehicle maintenance and basic engineering to work through the software, identifying those tasks of which they have experience which are then related to the appropriate competence elements. The system is self-explanatory and clearly plays a useful part in getting the candidates started on the APL process with a minimum of tutorial assistance.
- 12.6 The EGSPERT and ADEX programs, produced by the Further Education Unit in connection with their report *The Assessment of Prior Learning and Achievement: The Role of Expert Systems*, provide guidance to tutors assessing candidates in the APL process. The programs were developed, using expert system shell software, to be applicable to all subject areas. The example provided with the report is specific to electrical installation work, but it is understood that a version for motor vehicle maintenance has also been produced. The system is less user-friendly than 'Capability' and focuses on the underpinning knowledge rather than evidence of practical competence. It does, however, suggest ways in which computerised guidance can assist the assessor in making decisions on candidate competence.

- 12.7 The Computer Based Assessment System developed by Mast Learning Systems in connection with the Management Charter Initiative competences has some features in common with 'Capability', but is intended to be formative rather than to provide end point assessment. The candidate rates him or herself on each element and ratings are compared with those given by the supervisor as a basis for staff development and training. Because the system is based on self rating with little guidance other than the text of the performance criteria, it is unlikely that two candidates would apply the same standards to their ratings. Although undoubtedly a useful self-development tool, the system in its present form is insufficiently rigorous to be used as a basis for accreditation.
- 12.8 The Management Computer Based Assessment System is a specialised development of Mast Learning System's more general DNA program (Development Needs Analysis), which provides a mechanism for similar self-rating on a wide range of competences and attributes. Personnel departments and others may purchase this system either with a standard set of competences or with a facility to add competence elements to meet the needs of their own personnel.
- 12.9 'Guidelines', developed by Systems Applied Technology provides on-line help for women returners to identify their own skills and interests, and hence to receive guidance on jobs which might be suitable and the additional training which they require to meet the needs of those jobs. Although it has some similarities with the APL process, the main focus is on guidance and as a route to additional training. Software with rather similar aims was developed in the late 1980s by Waite Fell-Gordon Associates, with funding from the then Manpower Services Commission, but it is understood that this system is not operational.

#### On-going research

- 12.10 TECs have been invited to bid for funding to run APL projects, and two have been in touch with the project team for information about computer assisted assessment developments, but it is understood that these TEC projects are not intended to result in the production of any new software.
- 12.11. The developers of 'Capability' have been contracted by the Employment Department to produce a 'shell' version, which could then be further developed for individual subject areas. The existing Capability system is also to be trialled more extensively, in conjunction with Avon TEC.

#### Potential for further development

- 12.12 The Employment Department might wish to consider commissioning or distributing a package to provide an initial general explanation of the APL process as a preliminary for candidates who might then wish to move to a subject-specific package. An existing NCVQ video (not seen) might provide a suitable basis.

- 12.13 The 'Capability' package clearly has the potential for transfer to other subject areas and GES welcomes the move to develop a shell from which subject-specific packages could be produced. The competence elements for each NVQ could be transferred into the shell from existing files (for example from the NCVQ database). It is assumed that subject expertise would be needed to define the job roles and tasks which would be related to each element or group of elements. The developers of 'Capability' found that the text of the competence elements was not self-explanatory to intending candidates, and therefore developed the package so that the candidate approached the process via questions about his or her prior experience and the type of job/task which he had done.
- 12.14 Examples of acceptable evidence of competence for APL are given in the RSA guide on Accreditation of Prior Achievement. As noted in Section 11, it would be helpful for assessors and verifiers if a wide range of examples of acceptable evidence and performance could be made available for their guidance. Similar examples could be made available to APL candidates, who would then be helped to draw up suitable portfolios and to identify areas of competence or areas where additional tuition was needed. The examples would need to be specific for each NVQ. The mode of provision might be in booklet form, but examples with text or simple graphics might be accessed from the appropriate version of 'Capability'.

#### **Possible markets**

- 12.15 As the option to gain qualifications by the APL procedure is taken up more widely, the possible markets for computer assisted guidance in this area would extend to all centres and TECs.
- 12.16 Producing a 'Capability' shell should increase the number of subject areas for which it will be economical to provide customised software.

#### **Possible follow-up action**

- 12.17 It is suggested that the Employment Department should commission a project to build up examples of evidence of competence which could be made available to APL candidates, either in free-standing format or linked to a 'Capability' program. An outline project specification (G12.1) has been prepared.

## SECTION 13: ASSESSORS AND VERIFIERS

### Introduction

- 13.1 Two aspects of assessment and verification might be improved by using databases:
- the interpretation of evidence, including its sufficiency (this is covered in Section 11)
  - listings of assessors and verifiers.

### Main features

- 13.2 For many NVQs there will be a complex fabric of deliverers of education and training (especially colleges and other managing agents), employers who require the specific qualification for their staff, and bodies such as TECs and LECs which will have a major interest in encouraging and facilitating assessment. Awarding bodies, which will tend to have national coverage, will have the primary response for appointing verifiers and approving assessors, but will mainly be concerned with their ability to conduct the assessment/verification for different subjects. Because assessment will need to take place when the candidate is ready rather than in accordance with any awarding body created timetable, there could be considerable delays in making assessors available. Patterns may emerge in which TECs and LECs, and perhaps other local bodies, will wish to maintain lists of assessors and verifiers and be able to take action to make sure that a suitable assessor can be made available. While *verifiers* are likely to be directly appointed by the awarding body there is no inherent reason why a particular person should not be an assessor for more than one awarding body.
- 13.3 There therefore seems likely to be a demand for the creation of lists of assessors and verifiers and, subject to Data Protection legislation, such listing may often be provided on databases.
- 13.4 If a shell database structure and common standards could be agreed, this would enable different databases of assessors and verifiers to share information as appropriate, again subject to Data Protection legislation.

### Contribution to improving access, efficiency and effectiveness

- 13.5 The main gains in efficiency would be the ability to share data between databases in different areas and the removal of duplication in collecting and maintaining it. There could be improvements in access if the databases on assessors enabled assessment to take place when it was required and without any delays while assessors were identified.

### Current availability and use and on-going research

- 13.6 GES is not aware of any work being done in this context.
- 13.7 More general work on Coherence and Compatibility (including *Using Learning Information: A Guide to Handling Information in Education and Training Databases*) has been undertaken under the auspices of the Learning Systems and Access Branch and ADSET (the UK Association for Database Services in Education and Training).

### **Potential for further development**

- 13.8 It would be possible to build up a national listing of approved assessors. However, for this to be useful it would be necessary for there to be agreed procedures. An infrastructure would need to be thought through and developed.

### **Possible markets**

- 13.9 Awarding bodies, TECs and LECs might be interested in the database shell and in common standards for the creation of databases of assessors. These bodies would also be users of the data. So also might be centres, especially in relation to subjects where the centre did not have in-house assessors or needed to supplement them with additional ones from outside.

### **Possible follow-up action**

- 13.10 The Employment Department might commission a detailed feasibility study, including specifications for the procedures and infrastructure to make a linked database system work.
- 13.11 Following the feasibility study the Employment Department could also commission the development of shell database software.

## SECTION 14: FEEDBACK AND QUALITY

### Main features

- 14.1 Techniques described in this section involve the use of computers to manipulate statistics to provide evaluative summaries, and also the use of computers for reporting.
- 14.2 Uses of computers to provide feedback from awarding bodies to centres, and from centres to candidates, are described in Section 4 (Central Administration and Assessment and Section 5 (Local Administration of Assessment).
- 14.3 Computers can be used to aggregate and analyse data on candidate responses to individual questions and assessment tasks, whether or not the assessment is undertaken at the computer. Analysis of this type can provide a useful indication to the awarding body of the level of difficulty of individual questions and can also indicate areas of potential problems. For example, analysis of multiple choice questions will help the awarding body to identify questions which candidates are finding particularly difficult and distractors which appear to be unfairly misleading to otherwise competent candidates.
- 14.4 Computerised analysis of the relationship between assessment components can be used to highlight any mismatch between performance on the components, and hence to provide pointers to problems needing investigation. This may be at the level of comparing success rates for the whole population on two or three main components, or may highlight marked discrepancies at the individual level.
- 14.5 Computer reporting of numbers of candidates taking individual units could be used as an indicator of the relevance of these units to the total candidate population.
- 14.6 Monitoring candidate numbers and success rates by such factors as candidate age, gender, ethnic origin and mode of study, can provide information about the practical outworking of equal opportunities policies, and could provide pointers to hidden discriminating factors (for example, questions might be set in such a way as to disadvantage particular ethnic groups).
- 14.7 'Data capture' for all these techniques may be by inputting from paper forms but is increasingly likely to be by electronic data transfer from files held for other purposes. Optical mark readers are a fairly well established method for 'capturing' responses to multiple-choice questions or questionnaires and could be used for scanning for marks awarded by examiners of traditional question papers.

### Contribution to improving access, efficiency and effectiveness

- 14.8 The major contribution of computer analysis and feedback of this type, is to highlight areas where there appear to be problems and which should be further investigated. Computer analysis can thus contribute to the overall improvement of **quality** or (if no problem areas are identified) help to reinforce the credibility of the qualifications.
- 14.9 If the result of computer analysis is to identify evidence of apparent discrimination, for example against ethnic groups or in relation to the relevance of individual units, this could have an indirect effect on improving **access** to vocational qualifications.

### Current availability and use

- 14.10 Traditional item analysis has been used regularly by City and Guilds since its first introduction of multiple-choice questions in the late 1960s. Multiple-choice questions are not widely used at present by other awarding bodies. As far as is known, those who are considering venturing into objective questions are not carrying out detailed analysis or pre-testing.
- 14.11 More sophisticated item analysis based on Item Response Theory, has been used in the development of some general ability tests available in the UK but is not, as far as is known, used in connection with assessment for any vocational qualifications. A possible exception is the development of adaptive testing in numeracy currently being undertaken by City and Guilds.
- 14.12 In the early 1970s City and Guilds investigated the potential for analysis of marks awarded in response to traditional questions, but work was discontinued. As far as is known, no other awarding bodies currently analyse results on individual questions other than multiple-choice.
- 14.13 As far as is known, no work was been done on the marking or analysis of results for other forms of computer delivered testing such as fault-finding exercises, although System Applied Technology are aware of this as an area needing investigation.
- 14.14 City and Guilds' main system for processing examination results includes facilities to compare, for example, pass rates on individual components as a means of monitoring success rates and identifying possible problems in individual components. BTEC operates a 'moderating' instrument for some units as a means of identifying possible discrepancies in the grades awarded by centres. Some GCSE boards identify potential discrepancies in results by highlighting candidates whose grade in one subject differs by three or more grades from that in others. With the move towards locally devised assessment based on elements and performance criteria, some of these techniques will become irrelevant to vocational qualifications and new means may need to be developed to identify any potential problems.
- 14.15 Major awarding body computer systems have the potential to report on numbers of candidates for individual units, and City and Guilds reports these numbers in the statistics which it produces. It is not known whether this is used as a mechanism for feeding back to Lead Bodies information about possible unevenness in unit take-up.



- 14.16 The main vocational awarding bodies collect information about the age and gender of their candidates (with some small exceptions), but the extent to which this information is analysed or used as a basis for detecting potential problems is very variable. With the exception of SCOTVEC, no major awarding body collects information on the ethnic origins or disability of candidates, and most are reluctant to do so. However, it is envisaged that a second stage in the development of the National Vocational Qualifications Information System (NVQIS), currently being set up by the Statistical Services division of the Employment Department, will be to encourage awarding bodies to collect these data items. For many candidates, the information is already available at the centres.
- 14.17 The Department of Education and Science collects data on students enrolled in Further Education establishments annually via the Further Education Statistical Return (FESR). Information collected about each student includes course being followed, age, gender, mode of study, type of funding and (since November 1990) ethnic group. This potentially rich source of data is under-used and is not at present related to other data such as success rates or the composition of the population in the college's catchment area. It also has limitations in that it does not cover students or trainees preparing for vocational qualifications outside the FE system.

#### On-going research

- 14.18 The Employment Department is currently commissioning two projects, one to investigate the validity of the BTEC moderating instrument, and the other to investigate other potential methods of measuring performance variables as a way providing feedback and monitoring of results of centre-marked assessments.
- 14.19 The Employment Department is also commissioning work on possible means of investigating the validity of assessment.
- 14.20 The first phase of NVQIS will make use of data already collected by the awarding bodies, but analysis may reveal trends which have not previously been noticed. Subsequent stages of development will, it is hoped, lead to the collection of more data by the major awarding bodies and to systematic collection of data from the smaller awarding bodies whose data collection mechanisms at present are variable and incomplete. For FE colleges, comparison with enrolment data should also be possible.
- 14.21 In June 1991 DES issued a circular announcing its intention to require FE colleges to publish information about success rates in courses leading to qualifications. This will report numbers of qualifications gained in relation to numbers enrolled on the appropriate year of the course and to time taken to gain the qualification. However, this data will be published at centre level, by occupational area and NVQ level, and will not be aggregated or analysed at national level.
- 14.22 A project involving LTU, the Further Education Unit and a London borough is comparing the demographic analysis of college enrolments with that of the population of the borough in order to assess equity of provision.

## Potential for further development

- 14.23 Most item analysis used in educational assessment focuses on the analysis of multiple-choice questions rather than other formats, and is based on norm-referenced concepts, relating candidate performance on individual items to their score on the test as a whole. Item analysis of this type may continue to have a place in pin-pointing shortcomings in individual items, for example distractors which are not working properly. There is a need, however, to investigate the appropriateness for NVQs of other types of analysis; this should entail a literature search to discover what has been done in the context of other examinations, particularly in psychological testing and in the United States. The research should include analysis of question types other than multiple-choice and also the potential of item analysis based on the classification of those attempting the items into masters and non-masters.
- 14.24 It would also be valuable to investigate the potential for using Item Response Theory analysis in UK vocational qualifications, in order to improve the efficiency of the test in distinguishing between candidates near to the 'cut score' (pass mark). There is also a need to investigate techniques for scoring and analysing candidate performance on computer delivered assessments of fault-finding and other skills.
- 14.25 In the past, computerised analysis of multiple-choice questions has been in numerical form only. Current computer technology could provide an 'intelligent item analyser' which would give a verbal commentary on items needing investigation. For example, if an item had a facility value below a preset figure, the printout could say 'this item is very difficult; check that it is not ambiguous and that it tests knowledge which the candidate really should possess'.
- 14.26 More work could be done with data which is relatively easily available, on the relationship between success rates on individual assessment components and on the success rates of sub-groups of candidates. However, the NVQ system introduces difficulties for analysis which were not present in the old examination-based system. Because of the cost of assessment and of awarding body fees, it is quite possible that candidates will only be made known to the awarding body when they have fulfilled the requirements for an individual unit. Little or no information will be available to the awarding bodies about candidates who attempted a centre-devised assessment for the unit but were unsuccessful, or who registered for a course but failed to achieve any units.
- 14.27 Computerised systems could provide more informative feedback to candidates about their performance. This is likely to be most useful for candidates who were unsuccessful, and could be in the form of a 'weakness report'. Pitman Examinations Institute already provides limited information on weaknesses to its candidates for some secretarial examinations, but the practice could also be applied to the results of different sections of other externally marked written tests.

### Possible markets

- 14.28 Although the beneficiaries of some of the results of some of these developments might be individual centres or candidates, work would need to be undertaken primarily by the awarding bodies. It is therefore necessary to convince these bodies of the need for, and value of, statistical monitoring. Staff of some of the larger bodies are known to appreciate, for example, the potential of statistical monitoring of equal opportunities policies; some of the smaller and newer awarding bodies may have less appreciation of the issues.
- 14.29 It is suggested that progress should be made by means of pilot projects funded by the Employment Department and/or NCVQ, to investigate the potential of particular monitoring techniques. When the value of such techniques had been demonstrated, the requirement to adopt them could be gradually incorporated into the agreements for accreditation of NVQs.

### Follow-up action

- 14.30 No proposals are made for follow-up action in relation to national trends in take-up of qualifications and success rates as those trends will be monitored, at least in part, by the National Vocational Qualifications Information System, the DES requirement for colleges to publish success rates and (increasingly) through the monitoring of training credits.
- 14.31 It is suggested that the Employment Department should commission follow up projects; outline specifications have been drafted for the following:
- investigation of the applicability of alternative item analysis techniques and to UK vocational qualifications, specifically NVQs (G 14.1)
  - production of an 'intelligent item analysis' program, initially for conventional analysis (G 14.2)
  - survey and evaluation of existing and potential methods for providing candidates with more informative feedback (G 14.3).

## Part III: Conclusions and Recommendations

### SECTION 15: CONCLUSIONS OF PROJECT AND CRITERIA FOR ADOPTING CAA

#### Introduction

- 15.1 This section summarises the main conclusions of the project and suggests criteria which can be used by organisations considering the introduction of computer assisted assessment.
- 15.2 A major conclusion from the surveys and discussions undertaken during the project is that there are a number of concerns about the NVQ assessment system, as currently being implemented. Some of these concerns relate to the quality of the assessment, including the difficulties of providing assessment which adequately covers the range statements and special or rarely occurring situations. There are also concerns about the consistency of assessment based on the work of local assessors. For centres, the most serious concerns relate to the cost of assessment, the amount of assessor time involved, and the difficulties of handling the paperwork.
- 15.3 Relatively little use is currently made of computer assisted assessment within vocational qualifications; the major exceptions are in central and local record keeping. However, the investigation of CAA and CBT products has shown that there are a wide range of technological applications which could be employed in computer assisted assessment, and which could address some of the identified problems of the NVQ system.
- 15.4 It is notable that almost no use is made of computer delivered assessment in assessment for National Vocational Qualifications, despite predictions of the potential for using computers in assessment which have been made over the past 10 years. In part this may be due to the rapid changes both in computer technology and in qualifications which have discouraged development. It may, however, indicate that external funding is required if significant progress is to be made.
- 15.5 A striking finding of the project was the lack of contact between three groups of personnel, the awarding bodies responsible for vocational qualifications, psychologists concerned with general ability testing (sometimes in computer delivered form) and the producers of CBT software. Each of these three groups seems to be largely unaware of the potential relevance of other groups to its own work.
- 15.6 It is clear that computer assisted assessment could be extended and could offer considerable benefits to the NVQ system, although it should not be regarded as a substitute for observation of the candidate in workplace activities, and it cannot be expected to provide a complete solution to all the problems identified.
- 15.7 In view of the high level of concern amongst centres about the cost of assessment, it is suggested that further work should be undertaken to identify in more detail where the problems lie.

## **Types and characteristics of CAA applications**

**15.8** Applications of computer assisted assessment can be grouped into the following broad categories, as described in earlier sections of the report:

- a) Central and local record keeping, reporting and tracking of candidates (Sections 4 and 5).
- b) Computer aids to test production (Section 6).
- c) Computer delivery of assessments of skill or knowledge and of simulations (Sections 7 to 10).
- d) Help for assessors and verifiers, including help in decision-making (Sections 11 and 13).
- e) Help for APL candidates (Section 12).
- f) Quality and feedback mechanisms (Section 14).

**15.9** These applications differ in the level of technology required for their implementation, likely markets (local centres or national bodies), their applicability to different occupational areas and the nature of the contribution which they can make to the assessment process.

**15.10** The level of technology required for any particular CAA application can also be defined in broad categories:

- a) Applications requiring central computing power only - e.g. central record keeping, test production and feedback mechanisms.
- b) Applications requiring local availability of PCs with only the minimum of facilities - e.g. local record keeping, help for assessors and APL candidates and some computer delivered assessment.
- c) Applications requiring local PCs with slightly more advanced features, such as colour screens, faster operation and linked CD-ROM players, plus use of linear video - e.g. some computer delivered assessment and some special assessor training.
- d) Applications requiring the local availability of interactive video - e.g. computer delivered assessment.

**15.11** The complexity of the software and the cost of development will often increase in line with the complexity and cost of the hardware required to deliver the application. A principle of CAA development should be to use the minimum hardware configuration which will meet the requirements.

- 15.12 For some CAA applications, such as central recording and processing and support for assessors and verifiers, the market would be the awarding body or bodies responsible for the qualification. For local recording and probably for support for APL candidates the market would be the local centre. For aids to test production and for computer delivered assessments the market could either be the awarding body (delivering its externally set assessments to centres in computerised form) or the local centre (purchasing an assessment resource which would be used to contribute to the evidence of competence for each candidate).

#### Contribution of CAA to improving access, efficiency and effectiveness

- 15.13 The work undertaken for this project has shown that computer assisted assessment could play a valuable part in improving access, efficiency and effectiveness in the assessment for National Vocational Qualifications. The nature of the contribution which CAA could make will vary across the different CAA applications.
- 15.14 Computer delivery of assessment can make the greatest contribution to the effectiveness and quality of assessment. Nationally devised computer delivered assessments can provide benchmarks of standards and good assessment practice. They can extend the coverage and range of the assessment by testing candidates' underpinning knowledge and understanding and by providing assessment of rarely occurring or dangerous situations, as well as simulations of lengthy processes. Centrally devised computer delivered assessments, especially if marked objectively, will offer a greater consistency of assessment content and standards than can normally be obtained using locally devised assessment.
- 15.15 Other CAA applications which can play a role in enhancing the effectiveness of assessment are quality and feedback mechanisms and systems to provide better help and guidance for assessors and verifiers. Computer aids to test production will not necessarily improve the quality of assessment but should enable the quality to be maintained in conjunction with improvements in efficiency and access.
- 15.16 The efficiency of assessment can also be improved in the longer term by computer delivery and marking of assessment; these can free assessors to concentrate on those aspects of competence which must be assessed in the workplace. Central and local recording systems, which reduce the amount of paperwork, can improve the efficiency of the overall assessment process, although with no improvement in quality, as can computer aids to test production. Those CAA measures which improve efficiency should result in savings in assessor time and in the cost of assessment to the centre, and sometimes in candidate time, thus addressing the most pressing problems voiced by centres.
- 15.17 Introduction of computer assisted assessment will require initial investment. Improvements in long term cost-effectiveness in purely financial terms will be achieved most easily by CAA applications which are either inexpensive to buy or develop or which are relevant to large numbers of candidates.

- 15.18 Improvements in access to vocational qualifications can be achieved by measures such as the provision of computer-delivered and marked assessments (especially via open assessment centres) and better guidance for APL candidates, which reduce the reliance on local assessors. Measures to reduce the cost of assessment will also improve access by making qualifications more affordable for the individual candidate.
- 15.19 As has been noted earlier in the report, the benefits of CAA applications are inter-related.

#### Potential disadvantages of CAA

- 15.20 Introduction of CAA could in some instances bring disadvantages which need to be taken into account in assessing the usefulness of any particular initiative. In particular:
- If computer delivered assessment were the only available method this would inevitably restrict access for candidates for whom the necessary hardware was not available; open assessment centres and use of CAA as an alternative rather than essential method will reduce this disadvantage, as will a concentration on the minimum essential hardware requirement.
  - Introduction of CAA will inevitably incur costs. Those applications which are based on well-established technology and are relevant to large candidate numbers should produce compensatory savings within a time-scale of a few years; others may never be cost-effective in purely financial terms.
  - If computer delivered assessment were treated as a replacement for workplace observation rather than an enhancement of it, the overall effectiveness of the assessment would be reduced.

#### Criteria for adopting CAA

- 15.21 Organisations (centres, TECs, awarding bodies or Lead Bodies) considering adopting CAA will need to weigh the initial investment required against the expected improvements in quality of assessment, efficiency (including longer term cost-efficiency) and access. The cost/benefit ratio will vary according to whether the CAA application under consideration is applicable across all NVQs or is specific to one occupation or NVQ. For NVQ-specific applications the number of candidates involved will be an important consideration.
- 15.22 It is suggested that CAA applications should only be considered if they have the potential to offer at least one of the following:
- more effective assessment in terms of quality, consistency and breadth of coverage either without any long-term increase in cost or in circumstances where the importance of the improvement (e.g. in assessing emergency and rare situations) outweighs long-term increases in cost

- an improvement in the efficiency or cost-effectiveness of the assessment process (from the point of view of the candidates, the local centre or the awarding body) without any loss of effectiveness
  - improvements in access to assessment without any significant reduction in efficiency or effectiveness.
- 15.23 The major consideration in evaluating a proposal for the introduction of CAA should undoubtedly be the contribution which it can make to the assessment process rather than the innovative nature of the technology itself.
- 15.24 Any CAA applications for use within NVQs should also offer:
- overall compatibility with the philosophy of the Standards Programme and the NVQ Criteria
  - credibility to all parties, including candidates and the general public
  - user-friendliness
  - technology which can be introduced at acceptable cost (with particular reference to the cost to the centre)
  - long-term sustainability in terms both of cost and of the effort required to maintain quality of assessment.
- 15.25 Although development projects may be based on innovative and incompletely developed technology, it is essential that any proposed implementations of CAA on a wide scale are based on proven technology. It is therefore highly desirable that any proposals which rely on more advanced technology should be trialled on a small scale before



## ANNEX A

### TERMS OF REFERENCE

#### 1. Project Aim

- 1.1 The Contractor shall investigate the potential of Computer-assisted assessment (CAA) for improving access, efficiency and effectiveness of assessment for accreditation to nationally-recognised standards of competence and to inform the Department of possible interventions in the development and implementation of CAA.

#### 2. Project Objectives

2.1 The Contractor shall:

- i) review current, planned and future UK developments in CAA for accreditation to nationally-recognised standards of competence;
- ii) identify the potential market for CAA in the accreditation of nationally-recognised standards of competence in terms of:
  - a) the future delivery of NVQs/SVQs
  - b) the perceptions of need and future requirements in terms of systems and products
  - c) how such systems and products might be widely introduced and supported.
- iii) evaluate the potential of CAA with particular reference to increasing access, reducing cost, enhancing manageability and improving the effectiveness of assessment to nationally-recognised standards of competence
- iv) make recommendations on the possible roles for the Department in the development and implementation of CAA
- v) produce, as required, outline project specifications which reflect prioritised recommendations.

#### 3. Project Methodology

3.1 The general methodology proposed for the project shall involve:

- a) a technical study, to identify those applications of information technology which would be relevant to competence-based CAA;
- b) a market study, to establish which bodies likely to be responsible for assessment would find it economic to use which aspects of CAA.

3.2 The technical study shall involve a literature search; correspondence and questionnaires to identify relevant organisations and products; visits to producers of relevant products, and of computer-based training and other relevant IT products (including those which are not currently used for

assessment purposes but which could be so used); in-house testing of innovative programs.

3.3 The market study shall involve questionnaires and visits to users and potential users of relevant products including awarding bodies, Industry Lead Bodies, colleges, managing agents and a sample of large companies which might undertake their own assessment.

3.4 The studies shall be linked, and an iterative process is envisaged, covering:

- a) an initial survey of a wide range of existing and potential developments
- b) visits to a representative sample of producers and users
- c) visits to major organisations in the NVQ system (awarding bodies, ILBs)
- d) identification of existing CAA activity with potential for expansion to a wider range of users, including identification of the development need
- e) identification (in conjunction with the Department) of the technologies (including those not currently used for assessment purposes) which offer most potential for development
- f) further studies (including in-house testing) of these developments in more detail, with definitions of requirements to make them operational
- g) survey of, and consultation with, potential users of new developments (likely to be from awarding bodies, ILBs, colleges and training providers) to establish how far such developments would be taken up and used
- h) production of a set of prioritised recommendations for Department intervention, including an indication of the type and extent of intervention required and 'order of magnitude' figures of costs.

#### **4. Project Management**

4.1 The Contractor shall hold formal team meetings at least once a month, together with constant informal meetings, to ensure that work on the different strands of the project is meeting time schedules and is coordinated.

#### **5. Timescale**

5.1 The anticipated timescale is as follows:

APRIL

Literature survey, including information on current Department projects.

Search relevant advertisements e.g. CBT material and commercially produced candidate reporting systems.

Initial letters to awarding bodies, Industry Lead Bodies, some CBT producers.

Planning of initial visits; begin visits. Attend AETT International Conference (Developing and Measuring Competence) at Pontypridd to widen contacts.

#### MAY

Further visits to awarding bodies and ILBs. Further letters to producers of proprietary products; some in-house testing.

Identification of users of relevant systems (e.g. colleges and managing agents), marketing questionnaire and some visits.

#### JUNE

Continue visits to awarding bodies, ILBs, producers and users.

Discussions with the Department on developments to be followed up in more detail.

Begin visits related to developments identified for follow-up; in-house testing.

#### JULY

Complete visits for initial survey; in-house testing of products.

Continue follow-up work, including visits to, and other contacts with producers, users and potential users of developments not currently used for CAA.

#### AUGUST

Possible limited number of visits (e.g. in Scotland in the second part of the month) and to industrial organisations; more in-house testing.

Begin drafting report and project specifications for follow-up work.

#### SEPTEMBER

Final visits; in-house testing and follow-up.

Completion of final report.

### 6. Project Outputs

6.1 The Contractor shall submit to the Department a final report which shall contain:

- i) a review of current, planned and future UK activity and development in CAA in the assessment for accreditation of nationally-recognised standards of competence with reference to the roles identified in the project specification Introduction;

- ii) a review of current research relevant to the role of CAA in accreditation of nationally-recognised standards of competence in terms of the three roles identified for CAA;
  - iii) an identification of the potential market(s) for CAA in the future delivery of NVQs/SVQs, the needs of different market segments, and how such products might be widely introduced and supported;
  - iv) draft specifications for derivative projects which take into account organisational cultures; existing computer environments; the size of the occupational sectors and the nature of relevant standards; assessment procedures and the nature of acceptable evidence; local delivery infrastructures; user requirements; implementation requirements; product ownership; and distribution and support requirements;
  - v) a set of prioritised recommendations on the possible role(s) for the Department in the development and implementation of CAA;
  - vi) guidance which identifies the criteria and procedures for judging the cost effectiveness of CAA for nationally-recognised standards of competence.
- 6.2 The report shall include a review of technological developments which are not currently used for assessment for accreditation purposes, but which have the potential for development for this purpose.
- 6.3 The report shall be supported by descriptions of relevant products, which may be in the form of case studies or of promotional material for existing products.

## ANNEX B

### LIST OF VISITS AND DISCUSSIONS

#### **Conferences and Exhibitions**

AETT conference

TEED Open Learning conference

Human Resource Development exhibition

Learning Technologies Unit conference/exhibition

International Conference on Technology and Education, Toronto

Education, Training and Personnel Development Exhibition at NEC.

International Labour Office Conference, Tashkent

#### **Awarding bodies**

City and Guilds (two separate discussions)

Pitman Examinations Institute

LCCI and RSA (brief discussion in the course of visits connected with other projects)

Associated Examining Board

Royal Institute of Public Health and Hygiene

#### **Industry Lead Bodies**

Preliminary contact by telephone with Telecommunications Vocational Standards Council, Chartered Building Societies Institute and Hairdressing Training Board.

#### **TECs**

Visit to SOLOTEC (South London)

#### **Centres**

College of North East London

Wirral Metropolitan College

### **Software Producers**

Lasermedia

System Applied Technology Ltd

Lloyds Training Centre (included brief discussion with member of Banking ILB).

Hamlet Computing (producers of software for some GCE/GCSE Boards and professional bodies).

Selby-MillSmith

### **Miscellaneous**

Karen Moloney of Moloney & Gealy (consultants involved in standards setting)

LTU Learning Resource Centre

Astra Training Services

Institute of Manpower Studies

## ANNEX C

### Publications Consulted

**American Psychological Association**

'Guidelines for Computer-Based Tests and Interpretations',  
American Psychological Association, Washington DC  
(1986)

**Amner, P, King, K, Noble, P, Davies, R**

'A Brief Guide to Issues in Establishing Computer-Aided Records of Achievement',  
Project Report RP583  
Further Education Unit  
(1990)

**Araki, Y**

'A CAI and CMI system: considering student's achievement and consciousness',  
(unpublished paper)

**Araki, Y**

'CMI System for Evaluation of Learning Results',  
(unpublished paper)

**Bartram, D, and Marshall, L (Editors)**

'DRAT - Directory of Research into Automated (Psychological & Psychiatric)  
Testing',  
(draft for revised edition in preparation)

**Bartram, D**

'An Appraisal of the Case for the Adaptive Assessment of Knowledge and  
Understanding in the Delivery of Competence-based Qualifications', in Knowledge  
and Competence - Current Issues in Training and Education, (Editor H Black and  
A Wolf),  
Careers & Occupational Information Centre  
(1990)

**Bartram, D, Lindley, P A and Forster, J.**

'Review of Psychometric Tests for Assessment in Vocational Training',  
Department of Employment Group  
(1990)

**Bartram, D**

'Emerging Trends in Computer Assisted Assessment', in Personnel Selection and  
Assessment: Individual and Organisational Perspectives (Edited Shuler, H, Farr, J L  
and Smith, M),  
(1991)

**Bell, D**

'Exploring the role of Technology in Competence-based Assessment  
(unpublished paper)  
(1990)

**Business and Technician Education Council**

'Assessment and Recording Mechanisms for Competence-Based Specifications -  
Notes for Guidance 1991/92',  
Business & Technician Education Council  
(1991)

**Fennell, E (Editor)**

'Competence & Assessment, Compendium No 1',  
Department of Employment Group  
(1990)

**Fennell, E (Editor)**

'Competence & Assessment, Issue 14',  
Department of Employment Group  
(1990)

**Fennell, E (Editor)**

'Competence & Assessment, Issue 15',  
Department of Employment Group  
(1991)

**Fennell, E (Editor),**

'Development of Assessable Standards for National Certification',  
Department of Employment Group  
(1991)

**Further Education Unit**

'Interactive Video in Building Management Education', FEU/PICKUP Summary  
Report  
Further Education Unit  
(1987)

**Gitomer, D H, and Van Slyke, D.A**

'Error Analysis and Tutor Design', Machine-Mediated Learning, Vol. 2,  
Taylor & Francis and Mentor Systems, Inc  
(1988)

**Gitomer, D H**

'Individual Differences in Technical Troubleshooting', in Human Performance,  
Lawrence Erlbaum Associates Inc  
(1988)

**Gitomer, D H, and Yamamoto, K**

'Performance Modelling that Integrates Latent Trait and Class Theory', Journal of  
Educational Measurement Vol 28,  
(1991)

**Gregson, M**

'Credit Accumulation Project - Final Report',  
Haringey Education Services, Further Education Learning Quality Unit  
(1990)

**Hall, J**

'Interactive Video Learning Technology: A Training Aid for the Implementation of  
National Vocational Qualifications',  
Video Consultancy Service  
(May 1991)

**Howe, A**

'New technologies in training', Information Maps (Revised),  
Manpower Services Commission  
(1987)



**League for Innovation in the Community College**  
'Computerized Adaptive Testing: The State of the Art in Assessment at Three Community Colleges',  
League for Innovation in the Community College, Laguna Hills, California  
(1988)

**Lewis, C, and Sheehan, K**  
'Computerized Mastery Testing', Machine-Mediated Learning vol 2,  
Taylor & Francis and Mentor Systems, Inc  
(1988)

**Mansell, J**  
'The Assessment of Prior Learning and Achievement: The Role of Expert Systems',  
Further Education Unit  
(1990)

**National Council for Educational Technology**  
'Open Learning Focus', insert in OLS News Issue 37,  
National Council for Educational Technology  
(September 1991)

**National Council for Vocational Qualifications**  
'Assessment in National Vocational Qualifications, NCVQ Information Note 4',  
National Council for Vocational Qualifications  
(1988)

**Pitman Examinations Institute**  
'Business Administration Level I and Level II National Vocational Qualifications - Candidate Assignments and Trainer Resource Material',  
Pitman Examinations Institute  
(1991)

**Rediffusion Ltd**  
'Guide to selecting Technology-based Assessment Methods for NVQs',  
Employment Department Group  
(1990)

**Simosko, S**  
'Accreditation of Prior Learning, A Training Agency Perspective',  
Department of Employment Group  
(1990)

**SOLOTEC (South London Training and Enterprise Council)**  
'The SOLOTEC Training Account' (Information Pack),  
SOLOTEC  
(1991)

**Training Agency**  
'Guidance Note 1, A Code of Practice and a Development Model', Development of Assessable Standards for National Certification,  
Department of Employment Group  
(1989)

**Training Agency**

'Guidance Note 2, Developing Standards by Reference to Functions', Development of Assessable Standards for National Certification, Department of Employment Group (1989)

**Training Agency**

'Guidance Note 3, The Definition of Competences and Performance Criteria', Development of Assessable Standards for National Certification, Department of Employment Group (1988)

**Training Agency**

'Guidance Note 4, The Characteristics of Units of Competence', Development of Assessable Standards for National Certification, Department of Employment Group (1988)

**Training Agency**

'Guidance Note 5, Assessment of Competence', Development of Assessable Standards for National Certification, Department of Employment Group (1989)

**Training Agency**

'Guidance Note 6, Verification or Monitoring of Assessment Practices', Development of Assessable Standards for National Certification, Department of Employment Group (1989)

**Training Agency**

'Guidance Note 7, Project Management, Standards and NVQ/SVQ Development', Development of Assessable Standards for National Certification, Department of Employment Group (1991)

**Training Agency**

'Guidance Note 8, Consulting on Standards: Establishing the National Applicability of Draft Standards', Development of Assessable Standards for National Certification, Department of Employment Group (1991)

**Twining, J**

'Inside Information - A Door to the Future?', International Review of Education, Unesco Institute for Education (1986)

**Ward, W C**

'Systems for the Computerised Administration of Tests', (unpublished paper) (March 1990)

**Ward, W C**

'New Technologies and Occupational and Professional Testing', Educational Testing Service (1987)

**Ward, W C**  
'Planning for Computerized Testing',  
(unpublished paper)  
(1987)

**Ward, W C**  
'Measurement Issues in Computer-Based Assessment',  
(unpublished paper)  
(1988)

**Ward, W C**  
'Opportunities for the Use of Technology in Assessment',  
(unpublished paper)  
(1990)

'Accreditation of Prior Achievement - RSA Guide',  
RSA Examinations Board  
(1991)



## ANNEX D

### Results of Surveys

This annex shows the format of the questionnaires sent to centres and industry bodies and the frequency count of the replies received.

#### Centre Survey - Occupational Areas Mentioned (Question 2)

Many respondents mentioned more than one occupational area; others said 'all' or 'NVQs throughout the college' or (less helpfully) 'Level II'.

Five centres specifically mentioned SVQs. Where they also specified the occupational areas, they have been included in the following totals.

	No of centres
Business Administration	100
Information Technology	17
Accountancy	3
Retail	9
Care sector	18
Hairdressing and Beauty	24
Hotels and Catering	19
Agriculture and Horticulture	22
Horse Care	7
Small animal care	3
Engineering	15
Motor Vehicles	5
Construction	21
Trainers	4
Security	1
Publishing	1
Graphic Communication	1
Process Operations/Plastics Processing	2
Libraries	1
Photography	1
Management	1
Footwear	1
Bakery	1
Science	1
Steel	1
Dairy	1
Funeral	1

Guildford Educational Services Ltd

Computer Assisted Assessment: Centre Survey

Frequency Count

Total returns 250 (217)

Returns from education shown in brackets

1. Name of respondent . . . . . Position . . . . .
2. NVQ(s) with which concerned . . (see detailed list on previous page) . . . . .
3. Name of organisation . . . . .

Type of organisation (please circle)

College	YT/ET	Managing Agent	Company	Other (please specify)
217		29	11	3

CURRENT USE OF TECHNOLOGY IN ASSESSMENT

4. Please give the names of any software packages which you are currently using to:
  - a) monitor candidate progress on units and/or elements (e.g. 'Rocket', 'Standards')
  - b) deliver externally set knowledge tests by computer
  - c) deliver internally set knowledge tests by computer
  - d) test skills (e.g. keyboarding, software use) at the computer
  - e) test fault-finding/problem solving skills at the computer
  - f) other (please specify)

See detailed list on next page

Please complete the remainder of this questionnaire by ticking the relevant boxes. Any additional comments which you wish to make would be welcomed.

Current use of technology (Question 4)

a)	Software for monitoring candidate progress	
	Rocket	10
	Other	35
	11 centres mentioned that they were currently investigating available systems.	
b)	Externally set computer delivered knowledge tests	
	Hairdressing Training Board	3
	CG Inside Information	3
	Oxford Examiner	3
	Other	6
c)	Internally set computer delivered knowledge tests	
	Question Mark	6
	Word processing software	2
	Other	8
d)	Computer delivered skills tests	
	Keyboarding tutor programs (various)	23
	Standard software for word processing, database and spreadsheets	58
	Other	2
e)	Tests of fault finding or problem solving skills	17
f)	Other	25

Under e) and f) there were several references to software packages concerned with diet, nutrition and menu planning.

CARRYING OUT THE ASSESSMENT FOR NVQS & EVALUATING CANDIDATE EVIDENCE

5. Are you experiencing any problems in the demands which NVQ assessment makes on:

	Considerable Problems	Minor Problems	No Problems
Candidate time	70 (67)	93 (74)	49 (34)
Assessor time	132 (119)	66 (51)	19 (11)
Cost	103 (92)	60 (42)	41 (27)
Use of equipment	45 (45)	80 (72)	73 (46)
Other (please specify)	13 (12)	2 (2)	3 (2)
Additional comments . . . . .			

6. Are you experiencing any difficulty in arranging candidate assessment which:

	Considerable Difficulty	Minor Difficulty	No Difficulty
Covers all the elements in the standards	43 (39)	119 (98)	53 (41)
Covers all items in the 'range statements'	54 (49)	107 (88)	58 (44)
Covers situations which occur rarely in practice	61 (50)	98 (82)	47 (36)
Covers dangerous/emergency situations	54 (46)	81 (72)	72 (51)
Involves performance in the workplace	67 (63)	81 (67)	65 (47)
Other (please specify)	5 (5)	2 (2)	2 (1)
Additional comments . . . . .			

7. To what extent do you think there is scope for improvement in:

	Considerable scope for improvement	Some improvements needed	No significant improvements needed
The wording/clarity of the elements and performance criteria	66 (57)	105 (90)	47 (36)
The definition of the range of tasks/situations required	54 (47)	101 (87)	66 (52)
Guidance as to the difficulty or complexity of the tasks which candidates are expected to perform competently	83 (70)	87 (76)	55 (43)
Other (please specify)	7 (7)	3 (3)	1 (-)
Additional comments . . . . .			



8. Are you experiencing any difficulty in:

	Considerable Difficulty	Minor Difficulty	No Difficulty
Deciding whether individual candidates have fulfilled the requirements for a particular unit/element	18 (16)	107 (88)	92 (72)
Weighing the value of different types and combinations of evidence	27 (26)	110 (93)	78 (60)
Other (please specify)	1 (1)	3 (3)	2 (1)
Additional comments . . . . .			

ASSESSORS AND VERIFIERS

9. Are you experiencing any difficulty in:

	Considerable Difficulty	Minor Difficulty	No Difficulty
Finding enough suitably qualified assessors	46 (41)	49 (42)	116 (93)
Coordinating the standards of assessors within your organisation	24 (20)	89 (79)	94 (72)
Coordinating the standards of outside assessors e.g. work placement supervisors	99 (90)	71 (58)	35 (24)
Liaising with the awarding body's external verifiers/moderators	38 (34)	63 (55)	110 (85)
Other (please specify)	4 (4)	3 (1)	- (-)
Additional comments . . . . .			

RECORDING AND REPORTING OF ASSESSMENT RESULTS

10. To what extent do you think there is scope for improvement or additional help in the systems for:

	Considerable scope for improvement	Some improvements needed	No significant improvements needed
Recording candidate performance or assessment results	94 (90)	82 (69)	44 (24)
Monitoring progress of individuals/groups	73 (70)	103 (86)	44 (28)
Handling the associated paperwork	116 (105)	72 (61)	35 (19)
Other (please specify)	2 (2)	- (-)	- (-)
Additional comments . . . . .			

THE POTENTIAL FOR COMPUTERS IN ASSESSMENT

11. Which of the following items of hardware would be available for use in candidate assessment or recording/monitoring of results in the NVQs with which you are concerned? (please tick those available)

	Available now	Expected to be available within two years
<b>Computers:</b> IBM compatible microcomputer with floppy disc only	117 (96)	3 (3)
IBM compatible microcomputer with hard disc	187 (157)	7 (5)
BBC computer	69 (63)	2 (-)
Apple computer	47 (45)	7 (4)
<b>Peripherals:</b> Mouse	169 (147)	5 (4)
CD-ROM player	27 (26)	16 (12)
Interactive video equipment	38 (35)	19 (17)
DVI equipment	5 (5)	4 (2)

Other . . . . .

12. Which of the following assessment resources would you like to see made available for use in the assessment of NVQs in your area? (Please tick those which apply)

Banks of paper-based assessment/assignment material	160 (136)
Computerised tests of underpinning knowledge	181 (179)
Computerised tests of skills such as number, keyboarding, software use	123 (105)
Computerised tests of knowledge/understanding designed to assess students/trainees on entry to the course	177 (152)
Computer simulations (e.g. of fault-finding, or operation of complex plant)	108 (96)
Computerised expert systems to help evaluate candidates' evidence of competence	135 (117)
Computerised system for recording/reporting/monitoring candidates progress	194 (178)
Other computer/technology applications (please specify) . . . . .	4 (2)

13. Follow up: please give your telephone number if you would be willing to discuss these responses/ideas on the telephone with a member of the project team

Telephone . . . . .

Thank you for your help

ASSESSMENT FOR NATIONAL VOCATIONAL QUALIFICATIONS -  
 POTENTIAL FOR COMPUTER-ASSISTED ASSESSMENT  
 (ITO/LB questionnaire)

Frequency Count

Numbers of ITOs/LBs returning questionnaire forms: 55

Nil returns: 5

1. Name of respondent . . . . . Position . . . . .
2. Name of Organisation . . . . .

Please complete the questionnaire by ticking the relevant boxes, circling the appropriate response or filling in brief details, as appropriate. Any additional comments which you wish to make would be welcomed.

3. Type of organisation (please tick as many boxes as apply)

Industry Training Organisation	4	Lead Body	50
Sole awarding body	9	Partner in joint awarding body	31

CURRENT STAGE OF DEVELOPMENT OF STANDARDS

4. Please tick the appropriate boxes to indicate the current stage of development of the standards with which you are concerned, at each NVQ Level

	L E V E L			
	I	II	III	IV or above
Standards in place and being implemented (candidates already being assessed)	16	24	11	4
Standards prepared but not yet implemented	15	18	18	10
Standards in course of preparation	15	23	27	23
No standards drafting done so far	5	2	6	7

CARRYING OUT THE ASSESSMENT FOR NVQs

5. Which of the following do you see as posing problems or potential problems in the assessment of candidates for the NVQs with which you are concerned?

	Considerable Problems	Minor Problems	No Problems
Amount of candidate time required	7	30	11
Amount of assessor time required	28	19	5
Cost of assessment	23	22	5
Need for specialist equipment (please specify)	3	14	21
.....			
Arranging assessment which covers all the elements	14	27	9
Arranging assessment which covers the whole of the 'range statements'	20	24	10
Arranging assessment of situations which occur rarely in practice	19	26	7
Arranging assessment of dangerous/emergency situations	15	23	13
Arranging assessment for candidates who have not followed a conventional course or training programme	10	18	20
Other (please specify)	2	1	1
.....			

## ASSESSORS AND VERIFIERS

6. Which of the following do you see as problems or potential problems in relation to assessors and verifiers in the NVQs with which you are concerned?

	Considerable Problems	Minor Problems	No Problems
Finding enough suitably qualified assessors	19	18	15
Finding enough suitably qualified verifiers	17	17	15
Coordinating the standards of assessors	19	24	9
Monitoring the performance of centres	10	34	7

## CENTRAL PROCESSING OF RESULTS

(NB: This question is applicable only to organisations which are awarding bodies or principal partners in joint awarding bodies)

7. What problems are you experiencing in:

	Considerable Problems	Minor Problems	No Problems
Gaining comprehensive candidate data from centres	1	8	16
Processing results	2	4	19
Issuing certificates	1	6	17
Ensuring security of candidate data	1	4	19

POTENTIAL COMPUTER APPLICATIONS

8. Please indicate which of the following potential applications of computers you would like to develop for assessment in the NVQs with which you are concerned.

	Already have	Would like to develop	
		High priority (1-2 yrs)	Lower priority
Computer assistance in preparing assessment material	4	13	21
Computer-delivered tests of underpinning knowledge	4	11	25
Computer-delivered tests of skills (e.g. keyboarding, database use)	2	7	26
Computer simulations of:			
fault finding	6	8	18
plant operation	3	9	19
reaction to dangerous/emergency situations	1	15	14
Other (please specify)	-	1	2
.....			
Computerised guidance for Accreditation of Prior Learning	-	16	21
Computer assistance in decision-making in relation to candidate results	0	8	26
Computerised record keeping:			
at centre level	7	17	12
at national level	16	20	5

9. Follow up: please give your telephone number if you would be willing to discuss these responses/ideas on the telephone with a member of the project team

Telephone .....

*Thank you for your help*