



Qualifications and
Curriculum Authority



Llywodraeth Cynulliad Cymru
Welsh Assembly Government

Comparability study of assessment practice

Door supervision qualifications

October 2006

QCA/06/2710

Contents

Executive summary	1
PART ONE	6
1. Detailed centre visit findings	6
2. Overall analysis	22
3. Conclusions	26
4. Recommendations	26
PART TWO	28
5. A comparative evaluation of door supervision assessment materials	28
6. Test content and validity	29
7. Analyses of question difficulty and quality	34
8. Empirical analysis of perceived difficulty	42
9. Demands, difficulty and validity	47
10. Telephone interview findings	54
11. Overall strengths and weaknesses	59
12. Conclusions	60
13. Recommendations	62
Appendix 1: Acronyms	64
Appendix 2: The SIA test specification	65

Executive summary

Introduction

This comparability study is part of the regulatory authorities' ongoing programme of quality assurance monitoring of Vocational Qualifications (VQs). Under the terms of the Private Security Industry Act, 2001 all door supervisors are required to hold a licence issued by the Security Industry Authority (SIA). To obtain a licence, door supervisors must have an SIA recognised qualification. This study examined the consistency and quality of instruction and assessment practices associated with the National Certificate for Door Supervisors (Unit 1: Roles and Responsibilities for Door Supervisors, Unit 2: Conflict Management for Door Supervisors).

The awarding bodies offering this qualification are:

- Edexcel
- BIIAB/City & Guilds (BIIAB/C&G)
- National Open College Network (NOCN).

The outcomes of this study will be made available to the SIA and the awarding bodies offering the awards.

Methodology

The comparability study commenced in November 2005 and concluded in March 2006. A team of three scrutineers, including one team leader, each having expertise in vocational assessment and appropriate experience, was recruited to examine instruction and assessment practices across the college, employer and training provider centres approved to offer the qualification.

The team observed training and invigilation procedures as well as interviewing candidates and invigilators in 28 centres. Data collection was based on a common instrument provided by QCA.

In addition, a fourth scrutineer was recruited to carry out a comparison between the multiple-choice assessment methods used by each awarding body.

Six examination papers and methods of presentation were examined and compared in a series of ways according to a number of subjective and objective measures. An empirical exercise was

Comparability study of door supervision qualifications
carried out to estimate the relative difficulty of passing each of the Unit 1 tests, though this was not possible for Unit 2.

Centre visit findings

A judgement was made as to whether the awarding body requirements were being met by the centres visited. The team of scrutineers judged that the overall process was acceptable in 79 per cent of the centres visited and, therefore, not acceptable in 21 per cent of the centres. However, the details behind these judgements should also be taken into account.

Each participating centre was given prior notice of the scrutineers' visits, so the scrutineers anticipated full compliance with their awarding body's requirements. Scrutineers recorded what was observed on the day. In some cases it was clear that extra effort had been made for the visit but, even then, the standards fell short of awarding body requirements.

Strengths

The following were identified as strengths in some of the centres visited:

- many tutors had very good knowledge about the subject and supported the improvements in door supervisor practice
- good quality material was issued to the candidates
- good adherence to security and invigilation procedures (however, see associated weakness)
- adequate and, in many cases, very good facilities used
- good teaching/learning methods (however, see associated weakness).

Weaknesses

The following were identified as weaknesses in some of the centres visited:

- no recent visit by the awarding body representative
- limited variety of teaching methods used
- serious lack of examination paper security
- poor invigilation procedures applied
- variable level of help given to candidates with additional needs
- significant time delay for return of results and certificates in some cases

- incorrect, inconsistent and excessive information being taught and assessed
- some candidates perceived they were doing an inappropriate course, demonstrating a lack of understanding of the Private Security Industry Act (2001).

Test analysis findings

Analyses of the tests and the items considered the following features:

- content validity, item quality and its effect on test difficulty, empirical investigation of perceived item and test difficulty, the impact of cognitive and other demands on candidates.

In interviews representatives of the awarding bodies were asked about:

- question writing, test production, pass marks and item/test statistics, plans for future development.

Strengths

The following strengths were identified:

- good content validity
- no detectable difference in pass standards
- adequate writer qualification/training in all awarding bodies
- a commitment to developing item banking to improve quality and the consistency of standards.

Weaknesses

Some weaknesses were noted:

- the language of the questions was often judged too difficult for many candidates
- frequent language errors were made in some test papers.

Strengths and weaknesses

Some aspects showed both strengths and weaknesses:

- very different approaches to the challenge of testing inter-personal skills

- several different item types, some of which are good while others are inappropriate
- variation in the cognitive demands set by questions – good in one case, too low in another
- variation in the reading demands in different tests.

Conclusions

Adherence to the awarding body guidance, appropriate course presentation and suitable assessment practice was judged by the scrutineers to be effective in 79 per cent of the centres visited. While some good practice was observed in the quality of learning there were particular concerns raised regarding: invigilation and security of examination papers; confusion over resources and support arrangements for specific types of candidates, and assessment that does not always meet the learning syllabus. The compliance rate indicates that the majority of centres are delivering assessment to the required standards but that overall provision is inconsistent. Action taken to address the shortfalls identified would do much to improve the quality of the candidates' experience.

The awarding bodies differ in how they interpret the content demands of the test specifications. For Unit 1 it is not clear whether this affects content validity. Empirical investigation shows little evidence of differences in standard between the Unit 1 tests. The awarding bodies take very different approaches to the design of the Unit 2 tests. NOCN test mainly the learning of models and theories for appropriate behaviour. BIIAB/C&G try to test the application of learning fairly directly with tests based on videos of simulated incidents. Edexcel also base some of their test on scenarios, but express these verbally in quite long, written, question rubrics. A review should be carried out by, or including, an independent content expert, to consider whether these tests meet the SIA requirement to assess the *application* of skills in Unit 2. It should also consider whether the models and content being taught are sufficiently up to date. Several of the faults commonly described in textbooks and training materials for multiple choice item writers were common in these tests.

Recommendations

Many of the weaknesses noted in this study could be addressed with robust quality assurance by the awarding bodies, particularly visits to centres delivering the qualifications. Actions to be considered urgently:

By the awarding bodies:

- producing guidance and/or materials to encourage diverse methods for delivering content
- rectifying inaccuracies in course content
- ensuring robust invigilation and security arrangements
- working with the SIA to ensure that the qualification and assessment arrangements are suitable for each candidate's needs
- working with the SIA to reduce delays in post-assessment administration
- monitoring pass rates for Unit 2 by comparing these to Unit 1, rather than relying on random equivalence
- ensuring as soon as possible that items for Unit 1 tests are randomly chosen from an adequate bank of items within 'learning outcomes' or 'modules', so that their pass marks benefit from the assurance that 'random equivalence' brings
- moving quickly towards their declared aims of implementing fully featured item banking systems. This will improve the quality of the items that make up the tests and increase the confidence that certificate users may have in the standard of the qualification
- implementing systems for monitoring the quality of items, both at item review before they are used in tests and through analysis of real test data to identify items that need improving and delete faulty items.

By the SIA:

- it should relax the requirement for random item selection in Unit 2.

PART ONE

1. Detailed centre visit findings

1.1 Introduction

Under the terms of the Private Security Industry Act 2001, it is now illegal to work as a door supervisor without a licence issued by the Security Industry Authority (SIA). To obtain a licence, applicants must be aged over 18, pass an identity and criminal record check, and have an SIA-recognised qualification. The SIA has endorsed the following awarding bodies to offer recognised qualifications and approve trainers:

- BIIAB
- City & Guilds
- Edexcel
- NOCN.

BIIAB and City & Guilds (the 'Consortium') offer the Level 2 National Certificate for Door Supervisors. Edexcel offers the Level 2 BTEC Award in Door Supervision, and NOCN offers the Level 2 Award in Door Supervision.

The original members of the 'Consortium' were BIIAB, City & Guilds and NCFE. In April 2005, NCFE left the Consortium and formally applied to withdraw its new National Certificate for Door Supervisors from the National Qualifications Framework (NQF).

The accreditation start date for the Consortium qualifications was 1 December 2003, and for the Edexcel and NOCN qualifications 1 June 2004. The new door supervision qualifications have, therefore, been in the NQF for over a year. As these qualifications are an essential component of the SIA's licensing system, the regulatory authorities considered that a study of assessment practice associated with the qualifications was timely.

This report summarises the findings across 28 centres and will be made available to the Security Industry Authority (SIA), Skills For Security (SFS), which was formerly known as the Security Industry Training Organisation (SITO), and all the awarding bodies that offer these qualifications. Each awarding body is asked to respond in writing to the report, indicating how it intends to address any issues of concern highlighted by the study.

Purpose

The overall purpose of this study focussed on three main themes:

- to report on assessment practice for the new door supervision qualifications approved by the SIA for licensing purposes
- to note significant differences in assessment practice where found and to identify good practice
- to make recommendations for improvements in assessment practice where the outcomes of the study suggest these are necessary.

Objectives

For each of the new door supervision qualifications approved by the SIA, to investigate and report on:

- assessment requirements and implementation
- centre operations relating to quality of learning and candidate experience of assessment implementation.

Scope

The study will cover the new qualifications approved by the SIA for door supervision licensing, offered by SIA endorsed awarding bodies. Centres were selected at random across England and Wales on the basis of information supplied by the awarding bodies. Due to the time constraints and the size of the centre sample, the outcomes of this study should be taken as indicative. Nevertheless, the findings reveal aspects of delivery that require attention as well as notable examples of good practice that should be encouraged.

The comparability study scrutinised the awards from two perspectives:

- delivery of the assessment and aspects of the learning context
- analysis of the multiple-choice assessment papers.

This report is presented in two parts. The first addresses the findings from centre visits and the second addresses the evaluation of test papers.

1.2 Visit methodology

The team of scrutineers visited 28 approved centres across England and Wales (three in Wales). The visits took place between November 2005 and March 2006. The comparability study examined the consistency and quality of instruction and invigilation practices associated with the National Certificate for Door Supervisors (Units 1 and 2).

At each approved centre the scrutineers examined the instruction, assessment security and invigilation practices. They interviewed an average of five candidates and the key staff associated with the course in each centre.

The scrutineers were required to make judgements as to whether centres were adhering to the awarding body guidelines and requirements and to record their judgements on the data collection instrument provided by the regulatory authorities. In particular, examination paper security and invigilation were observed and judgements recorded.

1.3 Centre sample profile

The original intention was to construct a representative sample of centres to visit, matched against the number of centres approved by each awarding body and the spread across the different types of assessment centre. However, the composition of the final sample was dictated by the fact that many centres were not actually offering the courses when contact was made. The centres chosen were selected at random from the information supplied by the awarding bodies. Overall, 55 door supervisor (DS) centres were contacted and 28 visits made.

A further point of interest is that many of the centres in the sample were registered to offer a range of qualifications with more than one awarding body. This meant that additional time was required to confirm that a centre was offering the door supervisor qualification for the specified awarding body. The following table outlines the centre selection procedure.

Table 1. Approved centre selection procedure and numbers

Centre selection process	Number of centres
Original selection from lists supplied by awarding bodies	56
Centres that were not currently delivering the qualification	21
Centres not to be visited (visited recently by QCA on other matters)	1
Course cancelled	5
Other (could not make contact)	1
Centres visited	28

Centre type

The following table shows the types and numbers of approved assessment centres participating in the study.

Table 2. Approved centre type profile

Centre type	Number of centres
Training provider	18
Employer	9
College	1

Awarding body

The following table shows the number of centres listed by each awarding body. Many centres were registered with more than one awarding body and therefore it was difficult to judge the exact market share. An attempt was made to double the representation of the awarding body with smaller numbers.

Table 3. Approved centre numbers as supplied by the awarding bodies

Awarding body	Number of centres listed by the awarding body
BIIAB/C&G	177
NOCN	194
Edexcel	71

The following table shows the number of centres visited relating to each awarding body. The number of centres visited per awarding body very broadly represents the actual market share.

Table 4. Approved centre awarding body profile

Awarding body	Number of centres visited
BIIAB/C&G	8
NOCN	13
Edexcel	7

1.4 Programme delivery

The results of the observations made by the scrutineers in relation to the programme delivery are recorded in this section. At all times the scrutineers were making judgements based on the guidance supplied to centres by the awarding bodies. In most cases the results are expressed as a percentage of the number of awarding body requirements fulfilled by the centre. A result of 100 per cent would, therefore, imply the centres were judged to be fully complying with the guidance given.

Scrutineers' judgements were aggregated to produce the overall result in each table; these figures cannot be calculated by taking the mean of the results for each of the awarding bodies.

Delivery

Because of the countrywide delivery of the course, scrutineers collected information on whether the courses and assessments were being sub-contracted out. In all cases, the training was carried out by the registered centre. In many centres this involved 'buying in' the expertise, but none of the training courses was directly sub-contracted to another company.

Mode

Information was collected relating to the mode of course presentation.

All door supervisor centres visited delivered the qualification over four days. Some delivered the units in succession, but the majority split the course, offering the first unit in two days and the second unit over another two days during the following week. There were no differences between awarding bodies in this respect.

Table 5. Mode of delivery – four-day presentation

Awarding body	Four-day presentation mode %
BIIAB/C&G	100
NOCN	100
Edexcel	100
Overall use of presentation	100

Guided learning hours

The Learning and Skills Council define guided learning hours as:

'...all times when a member of staff is present to give specific guidance towards the learning aim being studied on a programme. This definition includes lectures, tutorials and supervised study in, for example, open learning centres and learning workshops. It also includes time spent by staff assessing a learner's achievements, for example in the assessment of competence for National Vocational Qualifications (NVQs). It does not include time spent by staff in the day-to-day marking of assignments or homework where the learner is not present. It does not include hours where supervision or assistance is of a general nature and is not specific to the study of the learners.'

Funding Guidance for Further Education in 2006/07

Door supervision courses are recommended to take 30 hours of guided learning time. The majority of centres allocated about 28 hours to the course – including the examination. In all cases this fell short of the recommended time. One BIIAB/C&G centre reduced the hours considerably because the candidates were deemed to be 'experienced'. One Edexcel centre running shorter days started the examination at 1.00pm on the final day, reducing the course time to about 20 hours.

Table 6. Guided learning hours

Awarding body	Centres offering recommended learning time %
BIIAB/C&G	88
NOCN	100
Edexcel	86
Overall	93

Location

The scrutineers collected data to record information about where the programme was delivered. The percentages relate to the number of courses presented at the registered centre, as opposed to other locations such as rented rooms or hotels.

Table 7. Delivery location in registered centre

Awarding body	Registered centres %
BIIAB/C&G	25
NOCN	46
Edexcel	29
Overall delivery in registered centres	36

Facilities

The quality of the facilities used ranged from purpose-built conference rooms to ad hoc, temporary locations. The facilities were checked to make sure they were adequate for running training courses and assessments. Overall, a very good standard was observed.

With the exception of one or two locations, all were considered adequate for their purpose. One NOCN centre was judged unacceptable because the facilities were very dirty, noisy and cold.

Table 8. Facilities for the training course

Awarding body	Centres judged to be providing adequate facilities %
BIIAB/C&G	100
NOCN	92
Edexcel	100
Overall	96

Delivery methods

Because of the intense nature of the short course leading to the examination, it was deemed appropriate that the scrutineers should observe and comment on the delivery methods used and feed this back to the awarding bodies. The delivery of learning was considered from the points of view of content and style.

The content presented for the vast majority of the courses was substantial. Comprehensive books and slides, prepared by the awarding bodies, were used in many cases. Almost without exception, the scrutineers commented favourably on the depth of knowledge and experience of the tutors. One unfortunate aspect, however, was that many centres took back the books they had issued when the course was complete. Prior to the courses, many centres encouraged self-study by the candidates.

The style of the delivery in many centres was equally encouraging, with group activities and interaction playing an important part. However, about half of the centres visited made excessive use of presentations delivered by overhead projectors. For such intensive courses this may appear to be the only way to complete the content. However, there were centres that broke up the presentations with activities, such as quizzes, to very good effect.

In many centres significant use of role play was observed and this proved to be an effective method.

At one Edexcel centre, the tutor made the mistake of criticising and contradicting the information on the SAFE slides, but did not make clear to the group which version they should learn for examinations. As noted later in this report, there are also issues about the accuracy and depth of material being presented in centres.

At one NOCN centre, the tutor missed out parts of the course because they were deemed inappropriate for the candidates' role.

Table 9. Delivery methods

Awarding body	Delivery methods judged to be appropriate %
BIIAB/C&G	100
NOCN	92
Edexcel	100
Overall	96

Specific guidance for assessment

There were a number of differences in the guidance supplied to centres by awarding bodies with respect to the invigilation and security procedures. The following table highlights these differences.

The scrutineers based their decisions on formal guidance provided. It is recommended that the awarding bodies consider the benefit of agreeing a common approach.

Table 10. Comparability of guidance

Awarding Body requirements	BIIAB/C&G	EDEXCEL	NOCN
Who can invigilate?	Should not be tutor	Not trainer	Trainer/Exam invigilator ie same person (1)
Should invigilators be trained?	'Suitably qualified and experienced staff'	Suitably experienced person should be 'Prepared for their invigilation duties by "Test Manager"'	'Suitably trained adult'
Paper delivery methods	Sealed envelope by post	By E-mail to nominated individual	Sealed envelope by post
Security arrangements	Opened in front of candidates. On examination completion Q & A sheets to be sealed in envelope within examination room	Opened in front of candidates. Paper copies to be stored securely. Email to be deleted	Opened in front of candidates. Post-examination – return by 'trackable method'
Paper return methods	By post within one working day	Q & A sheets into separate sealed envelopes	By post – 'timely completion and return'
Provision for additional needs candidates	25% extra time. Reader. Questions provided on audio tape	25% extra time. Prompter. (Any further	15% extra time. (or 25% extra time) (3) Reader/Scribe.

(learning difficulties)	An amanuensis (writer)	requirements refer to JCQ) (2)	
Provision for additional needs candidates (ESL)	Additional time (time not specified). Reader. English or bilingual dictionary (non-electronic)	25% extra time. Prompter. Bi-lingual dictionary. (Any further requirements apply to JQA) (2)	25% extra time. Bi-lingual translation dictionaries (non-electronic) if resident in UK less than 2 years
Procedures for readers and for writers	Separate room. If reader is also writer, extra invigilator required. Not tutor. Not relative	Separate room. Extra invigilator. Reader should not be teacher	Separate room. Extra invigilator. Reader should not be teacher or family member

1. NOCN/SITO guidance says that this is allowed, but it is not good practice.
2. JCQ – Joint Council for Qualifications (large document on its website).
3. NOCN Door Supervisor Guidance document states 15% extra time when reader is used. The NOCN/SITO guidance states 25% extra time when reader is used. This discrepancy must be dealt with.

1.5 Assessment process, security and invigilation

Pre-assessment security

The scrutineers observed:

- the procedures used for delivery of assessment materials to the assessment location
- the security arrangements used when the materials were opened/made available to invigilators/staff
- the time when materials were opened/made available to invigilators/staff.

BIIAB/C&G papers were delivered in sealed envelopes and in all the centres visited were kept securely until they were opened in front of the candidates at the start of the examination. This appeared to be a routine method that worked very well.

Edexcel email the questions to a nominated person. A set 'logging' procedure is then used to keep the papers secure after printing. Many centres seal the papers at this time and then open the envelope in front of the candidates. In all cases the scrutineers noted that the centres appeared to be complying with the requirements. However, it was noted that on two occasions the course tutor was the person responsible for safeguarding the papers. This is a conflict of interest and unacceptable practice.

At one NOCN centre the papers were kept in an unlocked drawer and had been opened in advance by mistake. This was a careless breach of compliance.

Table 11. Pre-assessment security

Awarding body	Compliance with awarding body guidance %
BIIAB/C&G	100
NOCN	92
Edexcel	100
Overall compliance	96

Invigilation

The invigilation procedures were a key part of the observations made by the scrutineers. The scrutineers recorded details on the level of centres' compliance with the specific awarding body requirements. These related to the process in general, but in particular:

- candidate identification (ID)
- number of invigilators, whether they were the course teachers, and any others present
- time-keeping arrangements
- information supplied to the candidates before start of examination (such as instructions related to fire, food, toilet, illness)
- arrangements for removing bags and notes from desks
- awarding body information made available to the invigilators.

In all cases the ratio of candidates to invigilators was acceptable.

Invigilation at all the BIIAB/C&G centres was judged to be meeting the guidance. In most cases it was carried out to a very high standard. A few small aspects were noted, for example, one or two centres did not specify a particular clock for the timing. In some cases the tutor acted to guide a less experienced invigilator. Checking of ID took a long time when done immediately before the examination. However, ID was always checked carefully.

At one Edexcel centre, a separate invigilator was used, but the person had no knowledge of invigilation procedures. The course teacher (who had also been responsible for the examination papers) explained all the procedures. At another centre, the invigilator voluntarily admitted to being untrained, but had developed a checklist from previous examination experience (which must also have been run incorrectly). No appropriate procedures were covered, the wrong times were used and no visible clock was used.

The invigilator at one NOCN centre conversed with several candidates during the examination to explain questions to them. As some candidates finished they left the room, but came back to shout messages (for example, about transport arrangements) to those still completing the paper. No procedures were explained and no clock was available. At a further two centres invigilators did not check ID. At one of these, the instructions supplied were vague and confusing. At a fourth centre, which did meet the invigilation requirements, the tutor/invigilator conversed with an English as a second language (ESL) candidate to explain the meaning of a question. The invigilator also kept a copy of the answer papers 'in case of discrepancies' to compare to the originals.

All NOCN invigilators explained the instructions. This complies with the awarding body guidance.

Table 12. Invigilation

Awarding body	Compliance with awarding body guidance %
BIIAB/C&G	100
NOCN	77
Edexcel	71
Overall compliance	82

Facilities for the examination

In most cases the same room used for instruction was rearranged and used for the examination. The scrutineers checked that the conditions were adequate (for example, in terms of room temperature, lighting and noise) and that the distance between desks was appropriate. However:

- at one Edexcel centre a window was open onto a noisy, floodlit sports pitch. The tables were in a 'U' configuration, therefore those on corners were too close together with the papers almost touching
- at two NOCN centres the tables were not rearranged adequately for the examination. Three candidates shared a small table in one case; four in another. The tables could have been rearranged easily.

This is unsatisfactory preparation of examination conditions and must be discouraged by the awarding bodies through guidance documentation and quality assurance visits.

Table 13. Examination facilities

Awarding body	Centres judged to be providing adequate facilities %
BIIAB/C&G	100
NOCN	77
Edexcel	86
Overall	87

Additional requirements

Arrangements for candidates with additional requirements varied considerably. Many centres were confused about what was allowed and what was not. In most cases an additional person was made available to assist and worked with candidates in a separate room.

Some technical issues arose in using a second room at BIIAB/C&G centres because of the video element, but most seemed to manage. In all cases noted during the visits, dyslexic candidates and ESL candidates were allowed 25 per cent extra time if this was requested. One centre did not allow ESL dictionaries, claiming that this is BIIAB/C&G policy. BIIAB/C&G guidance, however, says they are allowed.

A number of Edexcel centres were confused about how to deal with additional requirements and therefore a variety of methods were applied. In one case the course tutor acted as the reader, which is not permitted in the JCQ guidelines. Some centres were unsure whether interpreters were allowed. The information supplied by Edexcel to centres does not seem to give clear guidance on the procedures they should adopt.

An NOCN centre contacted the awarding body and asked whether dictionaries were allowed or whether extra time could be allocated and was told neither was permitted. NOCN guidance for assessment (in Appendix A) refers to the *NOCN administration handbook for centres* for other special assessment requirements. Some NOCN centres had made no arrangements for additional needs candidates. There is some contradictory information in the NOCN/SITO guidance (see Table 10).

Table 14. Candidates with additional requirements

Awarding body	Centres operating within awarding body guidelines %
BIIAB/C&G	100
NOCN	77
Edexcel	71
Overall	82

Marking and post-assessment security

The scrutineers judged the security of the papers after the examination to be in compliance with awarding body requirements. However, where papers were kept for use in the second unit examination, some level of trust was necessary, particularly where the tutor acts as the invigilator (NOCN).

In all cases, the completed papers are returned in sealed envelopes to the awarding body or a nominated company for marking. Where the awarding body required the same answer paper to be used for both parts of the examination, the answers to the first unit were kept securely and returned securely after the second examination.

Table 15. Marking returned to awarding body

Awarding body	Centres following awarding body guidance %
BIIAB/C&G	100
NOCN	100
Edexcel	100
Overall	100

1.6 Other issues observed

The scrutineers observed a range of issues that arose during their visits. Centres dealt with many of these minor issues competently. The following are examples of these issues.

During a BIIAB/C&G examination, one scrutineer examined paper 3935 and could have scored 7/10 or 8/10 on Part 1 and 10/10 on Part 2 without watching the video. The view of the team is that the video assessment method is valid, but with poor 'distractors' and only three options, the result is less reliable.

In an Edexcel centre, one candidate completed the examination in just 12 minutes.

At an NOCN centre, the staff had received the information pack only a few days before the visit, even though they had been running the course for several weeks. Coincidentally, the nominated person from NOCN was visiting on the same day as the scrutineers.

A small number of centres allowed candidates to miss out the training for Unit 1 if they produced evidence of previous training.

1.7 Awarding body results and appeals procedures

Results

The turnaround time for results and certificates was investigated.

BIIAB/C&G centres varied a little on the time taken to return results and certificates but, in general, it took one to three working days for results and one to two weeks for certificates to be returned. Centres were quite satisfied with this.

Edexcel centres tended to wait about a week for results and two to three weeks for certificates. Centres were generally satisfied with this.

For NOCN centres the messages were mixed. The awarding body states that results will be turned around in five days, but the average wait appeared to be one to two weeks for results and a further one to two weeks for certificates. Many centres expressed dissatisfaction with this and reported that they often had to phone and chase.

Table 16. Results and certificates turnaround time

Awarding body	Centres reporting acceptable turnaround time %
BIIAB/C&G	100
NOCN	31
Edexcel	100
Overall	68

Appeals procedure

There was little consistency regarding how candidates were made aware of the awarding body appeals procedure. There was also some confusion among centre staff as to whether one existed. The results in the following table show the outcomes. It is recommended that the awarding bodies make this clearer to centres.

In many centres, candidates were told that a procedure existed and that a note was pinned on the wall to provide further details if needed. This is good practice.

Table 17. Appeals procedure

Awarding body	Candidates made aware of procedure %
BIIAB/C&G	100
NOCN	62
Edexcel	86
Overall	79

Candidates

The team was asked to interview a representative number of candidates (usually five or six) at each of the approved centres visited. The total number of candidates interviewed was 151. The candidates were asked to give feedback on their general experience, course content and the assessment process. The following table shows the gender profile of candidates taking the course and examination.

Table 18. Candidate profile

Gender	Percentage of candidate cohort %
Male	96
Female	4

General experience

The candidates' perspectives of the course, in general, were very positive across all three awarding bodies.

A number of candidates expressed a desire for more time to be spent on restraint techniques as part of the course. The scrutineers' opinion, and that of most centres, was that this would not be a good idea as it would detract from the main purpose of the existing units. A further unit covering restraint and similar topics, however, could be considered for development.

Most candidates commented favourably on the quality of the instruction they had received. Some mentioned that being able to use bilingual dictionaries would have helped.

Several candidates (security guards, in most cases) suggested that they would never work 'on the doors' and had done this course only to get the status associated with holding a SIA licence.

Table 19. Candidates' general experience

Awarding body	Candidate satisfaction level %
BIIAB/C&G	100
NOCN	100
Edexcel	100
Overall satisfaction	100

Course content

Candidates' comments on course content and the time allocation were, in general, very favourable. Many suggested that it was, 'very useful, even for experienced door staff'.

Several candidates at NOCN and Edexcel centres thought that the presentation slides could have been written in simpler language and that the information on the slides was too detailed.

Table 20. Candidates' satisfaction with course content

Awarding body	Candidate satisfaction %
BIIAB/C&G	100
NOCN	92
Edexcel	100
Overall satisfaction	96

Assessment process

Candidates' perspectives on the assessment process and arrangements for invigilation were positive.

While a small number of candidates declared a dislike for the countdown clock on the BIIAB/C&G video stating that this, '...put them under unnecessary pressure', others said the same examination was the, '...best I have seen'.

Across all awarding bodies, several candidates commented that, '...the language used in the questions was confusing'. Several centres also considered that clearer English would be an improvement.

Table 21. Candidates' view of the assessment process

Awarding body	Candidate satisfaction %
BIIAB/C&G	100
NOCN	100
Edexcel	100
Overall satisfaction	100

2. Overall analysis

In general, although there is room for improvement in some areas, the situation is quite positive. Some centres were considered poor. However, many of the centres that were judged critically were attempting to do a good job. Some further guidance and, ideally, a visit from an awarding body representative would solve many of the problems noted by the scrutineers.

2.1 Overall results by awarding body and centre type

Awarding body

The following table shows the percentage of centres at which the scrutineers judged that the awarding body guidelines were being followed.

The results indicate that NOCN and Edexcel centres have some issues to resolve. These are highlighted in the main body of the report and in the 'Recommendations' sections of the report.

Table 22. Overall judgment on centres' compliance with guidance

Awarding body	Centres' compliance with awarding body guidance %
BIIAB/C&G	100
NOCN	69
Edexcel	71
Overall compliance	79

Centre type

The following table compares the percentage of centres that complied with their awarding body guidance, by centre type. The variation in numbers in the sample must be considered when comparing these compliance rates. Leaving the college results aside, the results show no significant difference in the rate of compliance between the types of centre.

Table 23. Centres following awarding body guidelines by centre type

Centre type	Centres following guidelines (number of centres) %
Training provider	77 (22)
Employer	80 (5)
College	100 (1)
Overall	79

2.2 Strengths

The following were identified as strengths in some of the centres visited:

Many tutors had very good knowledge of the subject and supported the improvements in door supervisor practice

Most trainers were very experienced and worked hard to improve the image and value of the DS role. Many had experience of the industry and were able to work well with the particular client group. In this respect, the conversion of the image of 'bouncers' to 'door supervisors' was being promulgated effectively.

Good quality material was issued to the candidates

Many centres had created good quality handouts to support the presentations. In general, the material supplied by the awarding bodies was good, but in some cases there were weaknesses involving incorrect information.

Good adherence to security and invigilation procedures (however, see associated weakness)

The majority of centres were doing a good job in this respect. In some cases where errors occurred this was due to ignorance on the part of centre staff rather than a wish to defraud or cheat in any way.

Adequate and, in many cases, very good facilities were used

This is identified as a strength because it was considered potentially problematic in the industry. Some training sessions were organised in nightclubs or similar venues. This was not considered appropriate by comparison with a specialist training room. However, it was ideal for role-play situations. Many sessions were organised in high-quality training rooms.

Good teaching/learning methods were demonstrated (however, see associated weakness)

On many occasions, the scrutineers observed very good use of role-play. This was an excellent teaching/learning method for the subject. Some tutors made the effort to post or email course notes and information in advance to candidates with English as a second language (ESL). At one centre, a competent female trainer put forward the female perspective for handling disputes as part of a very clear delivery. Most tutors emphasised key words for ESL and additional requirement candidates.

2.3 Weaknesses

The scrutineers also observed a number of weaknesses. Attention to these by the awarding bodies would improve the quality of assessment and overall provision.

No recent visit by the awarding body representative

Some centres have had no contact with a representative from the awarding body; many said they would welcome advice and the chance to ask questions. The scrutineers were told on many occasions that they were the first observers to visit the centre. Many of the issues raised in this report could be dealt with through awarding body visits.

Limited variety of teaching methods used

Many of the candidates did not have a frame of reference necessary to comment on the teaching methods used. However, there was room for improvement in many centres. The candidates (and scrutineers) commented on over-use of presentations using projectors and over-use of flipcharts.

In some sessions the slides used were difficult to see from the back of the room. The SIA/NOCN slide set is comprehensive, but it needs to be used in conjunction with other activities.

The 'company' perspective was pushed at times to the detriment of the course as a whole. For example, some tutors digressed from the syllabus by stating that the company policy should be sought in relation to drugs or prostitutes, rather than promoting ethical and legal standards.

Serious lack of examination paper security

In a few centres the course teacher had easy access to the examination questions prior to the course. This is bad practice and deemed totally unacceptable.

Poor invigilation procedures applied

Unqualified and inexperienced invigilators were among the key concerns that this study highlighted. In some cases, it was also clear that the 'so called' invigilator had been drafted in for the benefit of the scrutineer on that day. They were present to sign papers, deal with administrative issues and so forth as the tutor ran the examination sessions.

Common mistakes made during the assessment process included: times not noted, clock not used, incorrect times announced, tables not arranged appropriately. In most cases this was due to ignorance of the correct procedures. In some cases, the invigilation was done by the course teacher (even when the awarding body had specified that this should not happen).

Variable level of help given to candidates with additional needs

There appeared to be a general level of confusion over what extra help should be available for candidates with additional requirements, particularly those with English as a second language (ESL).

It appears that inconsistent guidance material has been issued by the awarding bodies in this respect. Awarding bodies should refer to the Good Practice Guide and note the sections relating to 'reasonable adjustments'.

A key issue raised by a number of sources was the quality of examination language and style. If this were improved, problems relating to ESL candidates and, indeed, other additional requirements issues, would be much reduced.

Time delay for return of results and certificates

The time lapse for the return of results and certificates varied considerably between and within awarding bodies. NOCN had particular problems in this respect.

Incorrect, inconsistent and excessive information being taught and assessed

Handbooks used by NOCN and Edexcel centres contained 1964 legislation for licensing and PEL drugs misuse. Clearly, this is superseded by legislation introduced in 2003 and that came into force in 2005. The NOCN examination also asked questions on this incorrect content.

A common handout included incorrect information concerning fire certificate requirements, according to a specialist tutor. It is noted that NOCN says that 'changes are in the pipeline'.

There was some inconsistent advice in material produced by (or on behalf of) the awarding bodies. One point of particular note was the guidance given on using a mobile phone near to a potential bomb. It would be more appropriate to advise that no-one should remain near a potential bomb in any circumstances.

The content of the material and slides used by NOCN and Edexcel was considered to be excessive in some areas. For example, issues associated with possible rape victims went beyond non-specialist police training.

Some candidates perceived that they were doing the wrong course

Many candidates, involved in college or other security roles for example, did this course as a route to getting the SIA licence. A significant number said that they had no intention of working as a door supervisor. It is noted however, that the candidates did not understand that the type of licence required is determined by the Private Security Industry Act, (2001). Nonetheless, both the SIA and

awarding bodies should consider the potential for flexibility in this qualification to support the needs of a diverse target market.

Comment

It should be noted that for every centre where a weakness was identified, a corresponding strength or at least an adequate performance could be found in another. This implies that the overall quality of provision is inconsistent and there is room for significant improvement.

3. Conclusions

While some good practice was observed in relation to the quality of learning, this report must focus on the assessment arrangements of the qualifications.

Adherence to the awarding body guidance, appropriate course presentation and suitable assessment practice was judged to be effective in 79 per cent of the centres visited. A number of key strengths were identified. Nevertheless, there were weaknesses in a number of areas.

Of particular concern is the invigilation and, in some cases, the security of examination papers. Awarding bodies must do more to ensure the integrity of assessment for these awards.

The level of compliance found by the scrutineers can be taken as indicating that the majority of centres are delivering assessment to the required standards, but that the overall quality of provision is inconsistent. The main areas of concern include:

- lack of direct contact/visits from the awarding bodies
- confusion over the resources needed by candidates with additional requirements
- delays in post-assessment administration
- assessment applied unfairly due to inappropriate, or even incorrect, learning content.

With attention to the weaknesses identified, the standard could be further improved in all centres.

4. Recommendations

The following recommendations are based on the weaknesses noted previously.

The scrutineers' findings across all centres visited indicate that these shortfalls apply to all awarding bodies. It is recommended that these issues are addressed by ensuring that the following actions are taken:

By the awarding bodies:

- awarding body representatives visit centres on a regular basis
- a variety of teaching methods is encouraged
- the need for examination paper security is stressed in guidance
- clear, consistent guidance relating to invigilation procedures is issued and its implementation monitored
- guidance in relation to candidates with additional needs is made clear and, where possible, consistent across awarding bodies
- awarding bodies make their appeals procedure clear to candidates and centres
- incorrect, inconsistent and excessive information is removed from syllabuses and handouts and, with immediate effect, is not assessed
- awarding bodies issue guidance to centres to ensure that candidates are on the most appropriate course.

PART TWO

5. A comparative evaluation of door supervision assessment materials

5.1 Introduction

These National Certificates are new qualifications. Although the awarding bodies concerned are very experienced in developing and running tests of a similar kind, any new qualification faces considerable difficulties. Among these are: ambiguities or lack of detail in the general specification that guides the tests; the need for new question-writing teams; the lack of model questions and test papers to emulate; and the absence of clear procedures for setting and maintaining standards. In particular, it takes time to build up the collection of good-quality test items needed if awarding bodies are to benefit from the opportunities that effective item banking can bring.

It is against this background that this evaluation of the measurement properties of the assessments of candidate door supervisors was carried out. The reviewer has considerable experience of test development, item writing and the academic study of test questions, over a period of more than 30 years, and it is therefore not surprising that the report contains many comments that are critical of awarding body practice. The intention, however, is to be positive and the criticisms should be taken as recommendations for improvements that the awarding bodies are well capable of implementing in order to raise the overall standard of assessment for these qualifications.

5.2 General methodology and materials

This part of the study addresses the comparability of the door supervisor (DS) qualifications in terms of the **quality**, **difficulty** and **cognitive demand** of the multiple-choice tests. In the first sections *quality* is assessed by looking at several indicators of good practice in writing items and constructing tests from them for qualification purposes. Following this is a report on an empirical exercise that estimated how difficult it would be for a candidate to answer each question correctly. This provides an estimate of the relative *difficulty* of passing the tests.

Cognitive demand is considered next. That is, the nature of the thought processes that are required of a candidate to answer the questions. This is examined to discover whether any of the tests demands more sophisticated thinking and mastery of the content than another. The next section reports the results of telephone interviews with representatives of the awarding bodies about the procedures followed to produce the items and tests and to ensure standards are maintained. Finally, the conclusions and recommendations resulting from all of the investigations are summarised.

Three awarding bodies currently set tests for this qualification: the National Open College Network (NOCN), Edexcel, and BIIAB/City & Guilds (BIIAB/C&G). One set of tests issued by each of them was studied, together with relevant documents.

NOCN: Unit 1 – Paper Ref: D/06
 Unit 2 – Paper Ref: C/D/09

Edexcel: Unit 1 – Test Number: DS-U1T6
 Unit 2 – Test Number: DS-U2T6

BIIAB/C&G: Unit 1 – Paper No: 3787
 Unit 2 – Paper No: 3934

Note:

The NOCN test papers carried copyright marks dated 2005; one of the BIIAB/C&G papers was dated as copyright August 2005; none of the Edexcel papers carried any date. All of the papers appear to have been compiled and administered during 2005.

6. Test content and validity

6.1 Specification coverage

The Security Industry Authority (SIA) lays down an outline specification that includes a 'Detailed Training Programme'. This specifies a two-part course:

- Roles and Responsibilities of Door Supervisors in the Security Industry Environment
- Communication Skills and Conflict Management.

Topics and sub-topics are listed for each part. The full list of these is attached as Appendix 2 to this report. Each of the three awarding bodies considered here assesses candidates against this specification in two units.

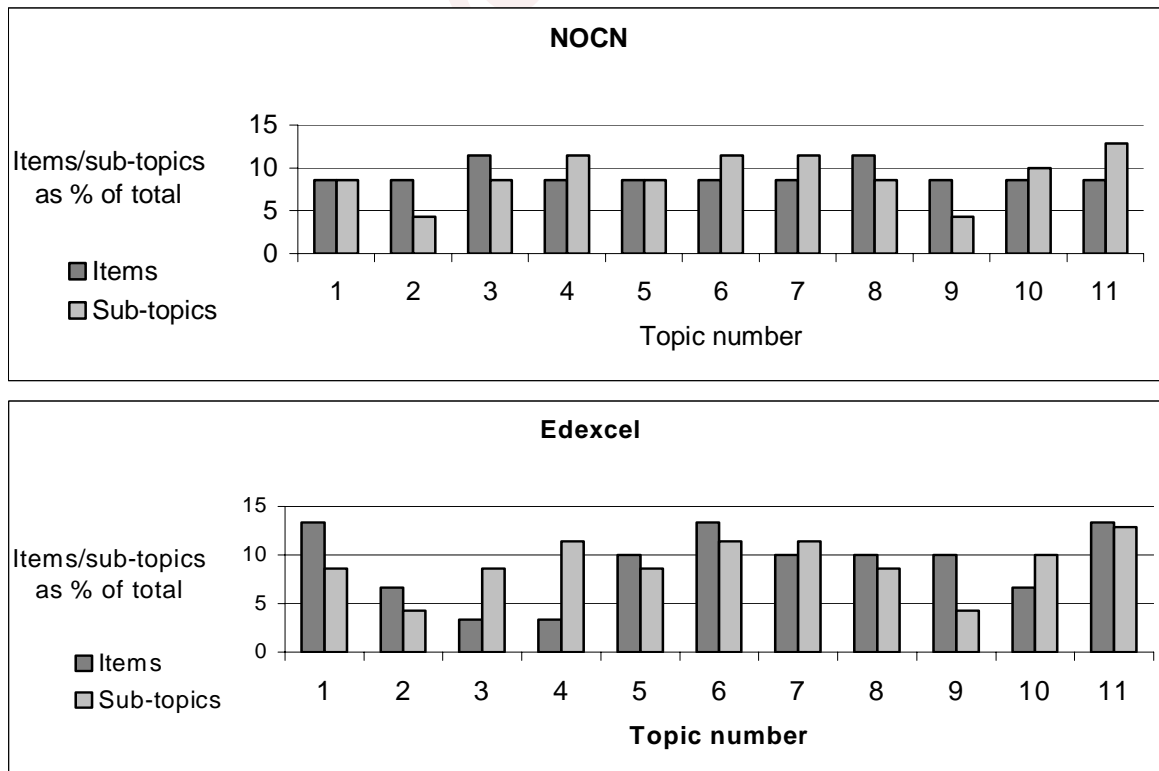
Unit 1 - Role and Responsibility of Door Supervisors

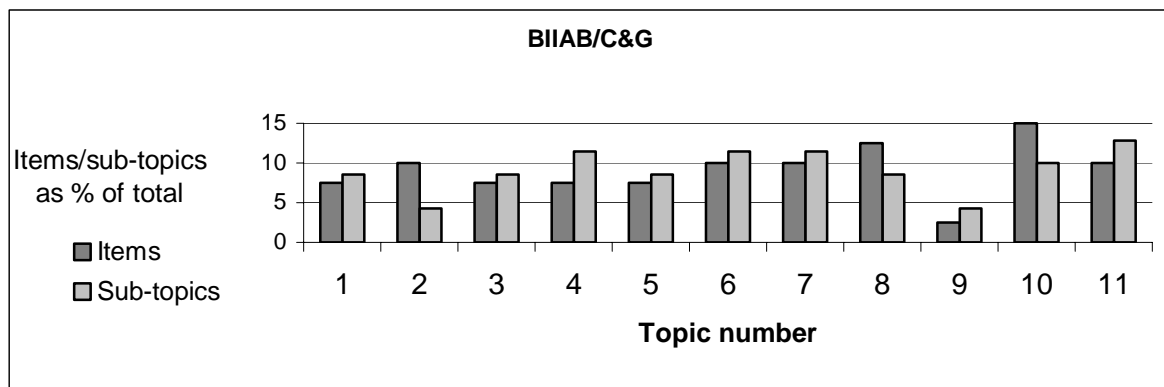
Eleven topics are listed (with the number of sub-topics or 'learning objectives'):

	<i>Topic</i>	<i>No. of Sub-Topics</i>
1	Introduction	6
2	Behavioural standards	3
3	Civil and criminal law	6
4	Searching	8
5	Arrest	6
6	Drugs awareness	8
7	Recording incidents and crime scene preservation	8
8	Licensing law	6
9	Equal opportunities	3
10	Health and safety at work	7
11	Emergency procedures	9

The three awarding bodies differ somewhat in how they convert this into a test specification. They vary in the number of items – there are 35 in NOCN's test, 30 in Edexcel's and 40 in BIIAB/C&G's – but this difference is probably not important in terms of coverage. Any test can contain only items from a sample of the 70 sub-topics and some sub-topics will deserve more items than others. The graphs below show how the items appear to be distributed across the topics. In each graph the dark columns show the percentage of test items that relate to each topic, while the light columns show, for comparison, the percentage of sub-topics in the SIA specification.

Figure 1: Graphs showing the tests' coverage of content





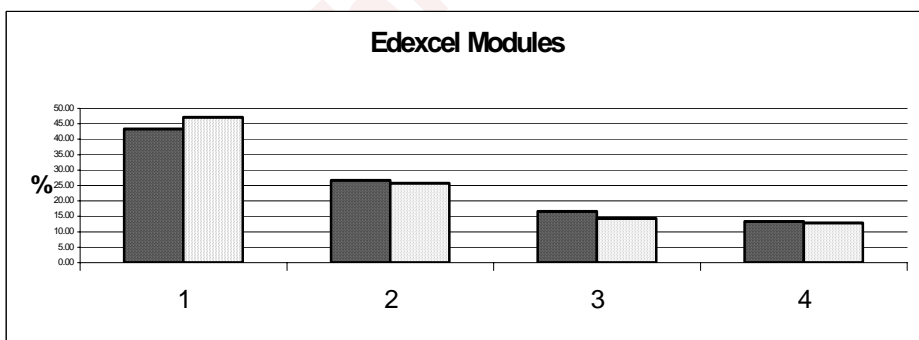
Interpretation

The awarding bodies seem to use different strategies. NOCN clearly sets three items for each topic and an extra item in each of the topics dealing with Law. BIIAB/C&G aims to set items in numbers roughly proportional to the number of sub-topics. Edexcel rearranges the SIA specification topics into four course modules as below.

Module	Sub-Topics
1 Behavioural standards, search and arrest procedures, and drugs awareness	33
2 Civil and criminal law	18
3 Equal opportunities and health and safety at work	10
4 Emergency procedures	9

The number of items matches the number of sub-topics in each module.

Figure 2: Edexcel module coverage of topics



Only a subject expert can decide whether these different strategies are equally valid. There is no apparent guidance to support any one over the others, and this would appear to be a matter for the regulatory authorities or SIA content specialists to determine.

All of the awarding bodies give equal importance to each topic or sub-topic. None of them makes any explicit judgement about the relative importance of topics or sub-topics, except perhaps in the small deviations from perfect proportionality that the graphs show.

Unit 2 - Communication Skills and Conflict Management

It is more difficult to compare the tests of Unit 2 of the qualification, since the awarding bodies adopt even more varied strategies and the tests cannot be compared as easily as for Unit 1.

The SIA specification divides this Unit into two sections (with associated sub-topics or 'learning objectives'):

<i>Section</i>	<i>No. of Sub-topics</i>
• Introduction to communication skills and conflict management	10
• Application of communication skills and conflict management	18

In addition, it lists three types of scenario in which the second section's skills might be developed:

<i>Scenario</i>	<i>No. of Sub-topics</i>
• Refusing entry to a customer	9
• Ejecting a customer from the venue	3
• Incidents inside the venue	7

NOCN sets 35 four-option multiple-choice items, which largely test knowledge and simple understanding of relevant vocabulary (eg *empathy*, *communication block*, *trigger*) or behaviour models (eg *POP*, *SAFER*, the '*Four As*'), and of strategies and explanations for behaviour. In the test studied they were distributed in these four categories:

<i>Category</i>	<i>No. of Sub-topics</i>
• Vocabulary	8
• Models	12
• Strategies	11
• Explanations	4

Edexcel once again rearranges the SIA content into four learning outcomes or modules:

- Customer care
- Communication skills
- Conflict management
- Situations leading to conflict

The test items are spread evenly across these modules. There are nine multiple-choice items (of which four are True/False, three are simple four-option and two are multiple-response, four-option items), followed by 16 four-option items based on four written narrative scenarios. The non-scenario items mostly address aspects of inter-personal communication (six items) and behaviour (three items). Thus, there is little testing of the models or vocabulary that are so prominent in the NOCN test.

The BIIAB/C&G test is entirely scenario-based, with the scenarios presented by video. There are four short recordings (each shown a total of three times), and 10 simple three-option, multiple-choice items that address issues that happen in, or are prompted by, each scenario. Although all of

the items are linked to the videos, they do not all depend on close watching: some of them test knowledge or understanding of taught elements and could be answered correctly without seeing the video. Even so, there is much less testing of vocabulary and models than in the NOCN test.

There is one major concern here. These three tests vary considerably in how they interpret the SIA specification:

- NOCN emphasises the learning of simple psychological theories and models to help door supervisors to behave appropriately.
- BIIAB/C&G tries to assess the skills for handling potential conflict and gives much less attention to direct testing of the knowledge underlying these skills.
- The Edexcel test lies between these two.

The SIA specification expects participants both 'to discuss communication skills and conflict management' and 'to observe, discuss and participate in scenario situations'. It does not require that the awarding bodies set questions in a scenario context, but it is difficult to see how a test can truly address a learning objective such as 'know how to eject a customer' without using simulation, role-play or scenario-based questions. Given that the first two of these are not practical in this context, BIIAB/C&G's video presentation seems, in principle, well suited to assessing these objectives. Of course, verbal narratives, as used in the Edexcel test, can also present scenarios effectively, though at the cost of increasing the reading load significantly. Only about three of the items in the NOCN test make any attempt to set a scenario, by describing a context in the stem of a single item.

A further concern has been expressed about the BIIAB/C&G video test. A knowledgeable candidate would be able to answer many of the questions without watching the video. Sometimes the video is simply irrelevant to choosing the correct response, as in:

Question: Why does the black door supervisor use 'Open PALMS'?

Answer: To signal non-aggression (C)

This, however, does not render the test invalid. The aim is to test the skills and knowledge needed to manage actual and potential conflicts, not to test watching and listening skills (as a language test might aim to do). The video is used to contextualise the assessment process and to reduce the amount of reading that would otherwise be needed to do this.

Some other items can be answered without seeing the video because the distractors are implausible. This will be addressed in Section 7.

6.2 Questions within topics

While considering the content of the tests one concern should be noted. The explicit models of verbal communication tested in the NOCN and Edexcel tests are very old-fashioned, based on theories from around the 1950s. Psychologists and linguists today view human communication as the construction of meaning by the reader/listener/watcher, rather than as the transmission of messages from sender to receiver.

7. Analyses of question difficulty and quality

7.1 Note on methodology

All of the analyses in Section 7 are subjective and a different judge would be likely to find different total numbers of cases of each problem in the items. The purpose of the analyses are, however, comparative, in that they should show conclusively which test contained most and which fewest examples of each feature judged. In order to ensure that a common standard was applied across tests the following procedure was followed. Analysis of one feature was completed before the next was begun. The first quarter (approximately) of the items in one test were analysed, then the first quarter of another, then the first quarter of the third. This pattern was repeated with the second, third and fourth quarters. The whole cycle was then repeated for another feature, taking a different test first each time. Usually, the early judgements were reviewed at the end of each analysis to ensure that the judge's standard had not changed in the course of the analysis.

7.2 Option plausibility

In a multiple-choice test the difficulty of questions is a function both of the *intrinsic* difficulty of generating/identifying the correct response to the question stem and the *interactive* effect of the other options. The intrinsic difficulty will be addressed more directly in Section 8. This sub-section looks at the role of the distractors.

Ideally, all of the options will be plausible answers to the question for candidates who are wholly ignorant of the test content, but only one will be plausible to a candidate who has mastered all of the content in the specification. If an option is not plausible to a candidate it cannot tempt them away from the right answer and thereby help measure their knowledge. For these analyses reviewers categorise weak options in three ways. (Note: these analyses are three variants of what is described in the literature as Nedelsky's Method.)

- The first, *Possible*, imagines a candidate who has some reading difficulty and so may not completely understand the question but can spot give-away clues linking the stem to the correct option. Or there can be obvious language errors that disqualify a distractor. Good questions should not have clues like these.
- The second, *Plausible*, forecasts the reasoning of a candidate whose reading ability is good, who knows nothing about the relevant content beyond the most basic general knowledge, but who can spot a 'silly' answer. The purpose is to indicate how many options the reviewers may discount on reading comprehension grounds alone, such as when an option does not address the question or gives a clearly non-legal answer to a legal question.
- In the third, *Likely*, the author of this review used his many years of experience in writing and scrutinising test questions and in working with question writers to eliminate unlikely options. These judgements were made before he had checked the marking schemes or any relevant content documentation, and in the few questions where he had picked up some specific information he ignored it. The purpose of this variant is to see how much test wiseness and general knowledge could raise someone's score above chance level.

The results of these analyses are converted into test scores as follows: the score on each item is the chance probability of getting it right after options considered 'impossible' have been removed. Thus, for example, the item score is 0.33 if one option is considered impossible in a four-option item, or 0.5 where two are. The results are reported in percentages and the ideal chance percentage is included as a baseline.

Table 24: Option plausibility and test score – tests

Test	Number of items	Ideal %	Possible %	Plausible %	Likely %
NOCN - 1	35	25	25	54	84
- 2	35	25	25	61	91
Edexcel - 1	30	33	35	55	80
- 2	25	29	31	54	87
BIIAB/C&G - 1	40	25	25	40	78
- 2	40	33	33	52	78

Table 25: Option plausibility and test score – qualifications

Test	Number of items	Ideal %	Possible %	Plausible %	Likely %
NOCN	70	25	25	58	88
Edexcel	55	31	33	55	83
BIIAB/C&G	80	29	29	46	78

Interpretation

There were a few options judged 'impossible' in the two Edexcel tests (four and three respectively). From a total of 192 options in the two tests this is not many – they only raise the 'possible' scores from 33 per cent to 35 per cent and from 29 per cent to 31 per cent – but there really should not be any impossible choices in a good test. There were none in the other tests.

A person able to read and understand the questions would be able to raise their score to 46 per cent on the BIIAB test, to 55 per cent on the Edexcel test and to 58 per cent on the NOCN test with no particular relevant knowledge. This means that the Edexcel and NOCN tests gave a greater number of inappropriate clues. This was largely done by including a greater number of options that were either implausible or obviously correct. For example:

Implausible

If an aggressive person invades your intimate space, which of the following might you feel?

D Safe and aware

Obvious

What is the essence of the Race Relations Act?

D It is unlawful to discriminate against any person on racial grounds

(cf 'because of their disabilities', 'because of their gender', 'because of their marital status')

Such a person would fail the Edexcel tests by about 4.5 marks in each test; would fail the NOCN tests by about 6 and 3.6 marks, and would fail the BIIAB/C&G tests by about 12 and 3 marks.

The pattern is similar for options judged 'likely' – the BIIAB/C&G questions contained fewer of the clues that a test-wise candidate could use to get correct answers without the relevant knowledge.

Based on these considerations:

- for Unit 1, 'plausibility' makes the Edexcel and NOCN tests significantly easier than the BIIAB/C&G test
- for Unit 2, 'plausibility' makes all three tests seem a little too easy for comfort, but they do not differ significantly in standard.

Much of the content being assessed is essentially common sense, though often presented in a somewhat formalised way. It is therefore reasonable to assume that an educated person, who can understand what they are being asked, should be able to answer quite a lot of the questions correctly without attending the course.

7.3 Length of options

One of the well-known faults in setting multiple-choice questions is to make the key option stand out by being a different length from – usually longer than – the distractors. The tests reviewed contained some items like this, and they were counted. The criteria used were:

- the key has more [fewer] words than all the distractors
- **and** the key extends to two lines [one line] with all distractors on one line [two lines]
- **or** the key is at least 50 per cent longer [33 per cent shorter] than the average distractor
- **and** the length differences reflect differences in syntactic complexity
- **and** there is no good justification for the difference (eg they are all titles of Acts).

The following example illustrates these criteria.

The spokesman gets agitated and begins to shout and swear, and is joined by several other males in the group, demanding entry. Which of the following would best allow you to deal with this situation?

- A You allow the whole group entry
- B You remain calm, and explain that if they do not change their behaviour you will continue to refuse them entry
- C You ignore the men and allow a different group of women entry
- D You lock the doors and walk away

The key, B, is the longest option. At 20 words it is two and a half times as long as the average of the others. A simplified phrasal group code† was used to indicate syntactic complexity. Here it gave:

- A: N.V.N.C
- B: N.V.C.V.A (N.V.N.N.V.V.N.N)
- C: N.V.N.V.N.A
- D: N.V.N.V.

showing that B was more complex, especially in containing five verbs and an adverbial conditional clause.

† V = verb phrase; N = noun phrase; A = adverbial/adjectival phrase; C = complement phrase. In some other questions additional codes were used, such as G for a negative. The exact form of this analysis was varied to fit the particular question but it was always consistent for the various options within a given question, which is what matters for the analysis in this section.

Table 26: The number and percentages of items with lengthy clues

Test	Items with lengthy clues	Percentage %	Number of items
NOCN	7	10	70
Edexcel	5	12.2	41 §
BIIAB/C&G	0	0	80

These clues will have made the NOCN and Edexcel tests rather easier than intended.

7.4 Reading difficulty

Potential reading difficulties were assessed at two levels:

The first level (average) is intended to represent Level 2 candidates. Questions judged problematic at this level were unnecessarily complex or linguistically faulty, in a way that would leave such a candidate uncertain about the intended meaning.

The second level (minimal) is more severe. It is intended to represent a minimally literate person. This level would include many non-native speakers of English who have reached the Council of Europe's 'threshold' level in English. Items were considered to fall in this category if they were judged to involve an excessive amount of reading, even if the words and structures were not particularly difficult.

Table 27: The number and percentages of items judged problematic at these levels

Test	Number of items	Reading – average level	Reading – minimal level	Reading – average level %	Reading – minimal level %
NOCN	70	6	19	8.6	27.1
Edexcel	55	7	11	12.7	20.0
BIIAB/C&G	80	2	9	2.5	11.3

Interpretation

About one quarter of NOCN's questions, a fifth of Edexcel's and a ninth of BIIAB/C&G's were judged to be problematic for minimally competent readers. This represents a serious literacy problem, especially for candidates who have not been educated primarily in English to GCSE standard. Several questions, especially in the Edexcel test, contained language that would be unreasonably difficult for most candidates.

§ not including the 14 items whose options are: A True B False.

In the NOCN test the main source of these problems was complex patterning of the options, as shown in this example:

Q11 Which of the following statements correctly outlines the attitude and behaviour cycle?

- A My behaviour affects my attitude which affects your behaviour which affects your attitude
- B My attitude affects your attitude which affects my behaviour which affects your behaviour
- C My attitude affects my behaviour which affects your attitude which affects your behaviour
- D My behaviour affects your behaviour which affects my attitude which affects your attitude

The semantic density of these options is very high with 11 semantic units in each option; weak readers will have considerable trouble coping with this. Also, the syntactic structure of each option consists of a main clause followed by a relative clause, which is then followed by another relative clause. Although native speakers will not be troubled by these linear relative clauses they are, in fact, ambiguous in that the reference of each *which* could be the preceding word, or the whole sentence up to the preceding word.

There are 24 possible combinations of the four key phrases – my/your attitude/behaviour – of which 12 start with 'My'. Finding one of these 12 from memory would be unreasonable, and candidates must analyse them to exclude those that are wrong. There is nothing cyclical about any of them, despite the stem referring to a *cycle*.

In the Edexcel tests, reading difficulty was more often due to the amount of writing used to present a scenario to the candidates. The most extreme case used 41 words in the initial description of the scenario, but still needed a further 117 words to present the stems of the four items based on it.

In the BIIAB/C&G tests, the high level of reading difficulty mostly arose from the use of vocabulary that would be unfamiliar to weak readers (eg *in a proactive way*) or, in the case of the video scenarios, by asking questions about things that didn't actually happen. Answering correctly requires an understanding of counter-factual grammatical structures (eg *If the customer had shown signs of mental illness, how should the door supervisor have responded?*). These structures are notoriously difficult for non-native speakers of English to process successfully.

7.5 Text highlighting

In all forms of assessment the questions are meant to convey to candidates, as simply as possible, the task they are meant to show they can carry out. The question should not 'get in the way'. Candidates can be helped to understand the question by making its language as natural as

possible, but there are other strategies that can also help – the simplest is to use either a bold or an italic font to highlight key words.

In these tests NOCN makes this difficult by printing every question in bold, and never using italics. The following words and phrases are capitalised:

- negatives: LESS, NOT (four out of five times),
- emphases: MAIN, USE OF FORCE, EVERYONE.

Because of the nature of these tests, especially in Unit 1, there are many technical words which are printed variously in single or double quotation marks.

In the test papers scrutinised, Edexcel prints in a regular font except for the texts that set up each scenario, which are in bold. No other form of highlighting is used, although *not* and *never* do each appear in one item stem.

BIIAB/C&G prints in a regular font and uses single quotation marks several times to highlight technical or quoted words. The following are shown in bold:

- negatives: **never, not;**
- emphases: **most** (twice), **first, must, main** (three times), **greatest** (twice), **best.**

The BIIAB/C&G strategy of highlighting key words helps focus candidates' minds on the key feature of the question.

7.6 Other aspects of question quality

a b c d distribution

Good practice requires test constructors to ensure that each option – a, b, c and d, or t and f – is used as the key more or less equally often, so that candidates will not use, or attempt to use, inappropriate guessing strategies. In these tests the key was distributed as follows:

Table 28: Distribution of key response across options

Four- and three-option items:	a	b	c	d	Total number of items
NOCN	15	20	21	14	70
Edexcel	13	4	13	11	41
BIIAB/C&G: 1	11	11	9	9	40
BIIAB/C&G: 2	13	15	12	-	40
Two-option items:	t	f			
Edexcel	9	5	-	-	14

The only serious observation to note here is the shortage of 'b' keys in the Edexcel tests: in fact, there were no 'b's at all in the 20 four-option items in the Unit 1 test. Such an extreme case of uneven option use is likely to be noticed by some candidates, especially if they have much experience of multiple-choice tests, and would puzzle or worry those who noticed it.

Sex stereotyping

One specific feature of the BIIAB/C&G Unit 1 test caused concern. Within its 40 items, there were 12 in which a door supervisor was stereotypically cast as a male, as in:

*What should a door supervisor be **most** concerned about?*

C Imposing his authority

This can be overcome and there are several strategies used by the other awarding bodies to avoid this problem. Most common is to address the candidate in the second person ('*Who would **you** attend to first?*'). An alternative is to ask about door supervisors in the plural, so that 'they' is the appropriate form for the answer ('*What is the primary reason for it to be important that door supervisors are able to...*', '*So that they can inform the police.*') Thirdly, it is becoming increasingly acceptable to use 'they' as a singular pronoun of indeterminate gender ('... **a person** . . . *refuses to be searched?*' '*Refuse **them** access*').

Language errors

A number of items contained language errors. These varied from simple typographical slips, such as *non verbal* for *non-verbal*, to the frequent misuse or omission of apostrophes and question marks, to printing *port* instead of *porter* (this makes a serious difference), to grammatical errors that made one or more options ineligible as answers, as in the example:

Guests will respond to requests from a door supervisor more effectively if

A Uses a loud voice and move the person with reasonable force

This option has no subject for its main verb, and the second verb appears to be an imperative form; these inaccuracies are sufficient to make almost any candidate balk.

The number of errors in each test was counted.

Table 29: Number and percentages of items with language errors

Test	Items with language errors	Percentage %
NOCN	3	4.3
Edexcel	14	25.5
BIIAB/C&G	6	7.5

Errors in spelling, punctuation and grammar at best convey a bad impression of the awarding body involved. At worst they will influence candidates' selection of options, usually with the effect of making the question easier than intended. Item vetting and proofreading should catch most problems of these types. There should be no such errors in the final tests.

The Edexcel tests had many errors, ranging from minor to serious.

NOCN's errors were all typographical, with an extra word left in or a simple one omitted.

BIIAB/C&G's errors mostly involved treating *premises* as a singular noun, as in '*When a licensed premises is open*'. It would be more straightforward to use the approved form '*When licensed premises are open*'. In The Licensing Act (2003) *premises* is never a singular noun.

8. Empirical analysis of perceived difficulty

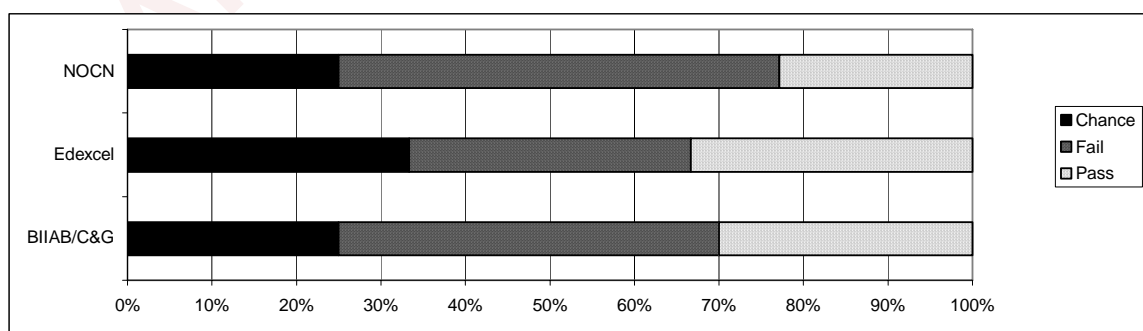
8.1 Pass marks of multiple-choice tests, and chance

It should be borne in mind that there is a base-line, the 'guessing level', which any candidate could be expected to score by chance alone. These tests use only two- and four-option items, with a variable chance of one-in-two or one-in-four of being correct by pure luck.

Unit 1

The graph below shows the amount of knowledge, in percentage terms, candidates must demonstrate to pass.

Figure 3: Amount of knowledge required to pass: Unit 1



The black strips show the chance effect; this is greater for the Edexcel test because 10 of its 30 items are two-option items. The grey strips are the important ones: they show the amount of knowledge, in addition to the chance score, that is demanded from candidates. The Edexcel pass mark, of 33.3 per cent + 33.3 per cent, will be less safe – even if the questions were more difficult –

than the BIIAB/C&G pass mark of 25 per cent + 45 per cent, or NOCN's of 25 per cent + 52 per cent.

In raw mark terms, beyond the chance score:

Edexcel	demands	10	extra marks
NOCN	demands	18.25	extra marks
BIIAB/C&G	demands	18	extra marks

The absolute binomial probabilities of someone passing purely by chance – with absolutely no knowledge at all – are (rounded):

Edexcel	1 in 600 thousand
NOCN	1 in 7 billion
BIIAB/C&G	1 in 352 million.

showing again that candidates are more likely to pass the Edexcel test by luck, and less likely to pass the NOCN test by luck.

More realistically, consider two candidates who definitely know 50 per cent and 40 per cent respectively of the items, and guess the rest randomly. What chance will they have of passing?

Table 30: Probability of passing with low levels of knowledge – Unit 1

% Known	Chance of passing (%)		
	NOCN	Edexcel	BIIAB/C&G
50	0.8	38.2	10.2
40	0.03	1.9	0.7

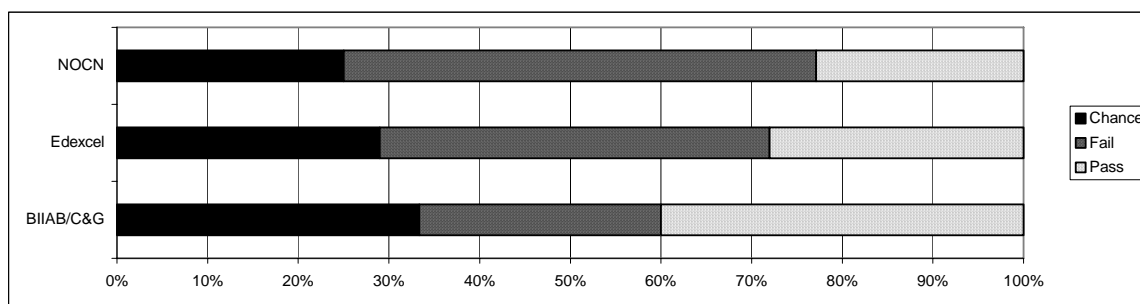
Knowing half the items, an Edexcel candidate is 38 per cent likely to pass, compared to having just a 10 per cent chance in the BIIAB/C&G test and less than 1 per cent chance in the NOCN test. The candidate who knows just 40 per cent still has a 2 per cent chance of passing with Edexcel, but less than a 1 per cent chance in the BIIAB/C&G test and virtually no chance in the NOCN test. (Remember, too, that NOCN's hurdles within modules make it more difficult than this to pass that test.)

This calculation suggests the Edexcel pass standard should be raised from 21 to 23 to make it comparable to BIIAB/C&G and to 25 to make it equivalent to NOCN, in terms of the effects of chance.

However, only the role of random guessing has been considered so far; the test standards cannot be compared without estimating the difficulty of the items, and this is considered in Section 8.2.

Unit 2

Figure 4: Amount of knowledge required to pass – Unit 2



The graph above shows similar calculations for Unit 2 and the effects of chance suggest that the BIIAB/C&G pass mark, at **33.3 per cent + 26.7 per cent**, will be less safe – even if the questions were more difficult – than the Edexcel pass mark at **29 per cent + 39 per cent**†, or NOCN's of **25 per cent + 52 per cent**.

In raw mark terms, beyond the chance score:

Edexcel	demands	9.75	extra marks
NOCN	demands	18.25	extra marks
BIIAB/C&G	demands	10.67	extra marks

Considering again two candidates who definitely know 50 per cent and 40 per cent respectively of the items, and guess the rest randomly, what chance will each have of passing?

Table 31: Probability of passing with low levels of knowledge: Unit 2

% Known	Chance of passing (%)		
	NOCN	Edexcel	BIIAB/C&G
50	0.9	39.6	94.0
40	0.04	11.3	57.6

Knowing half the items, a BIIAB/C&G candidate is almost certain to pass by random guessing, compared to having just a 40 per cent chance in the Edexcel test and less than 1 per cent chance in the NOCN test. The one who knows just 40 per cent still has a 58 per cent chance of passing BIIAB/C&G, an 11 per cent chance in the Edexcel test and virtually no chance in the NOCN test.

This calculation suggests the BIIAB/C&G pass standard should be raised from 24 to 27 to make it comparable to Edexcel and to 32 to make it equivalent to NOCN, in terms of the effects of chance.

† The odd figures for the Edexcel test are calculated by combining the chance effects for the two-option and four-option items.

However, as noted previously, only the role of random guessing has been considered so far; the test standards cannot be compared without estimating the difficulty of the items. Unfortunately, the method described for Unit 1 in the next section cannot be applied to Unit 2, because of the considerable variation in formats used for testing Conflict Management. This problem will be addressed in Section 9.

8.2 Investigation of question concept difficulty (Unit 1)

In most empirical studies of comparability, actual candidate performances are compared. For multiple-choice tests this requires reasonably large samples of candidates (several hundreds) to attempt at least two of the tests being studied, or their own test together with a highly reliable reference test. Neither approach was possible here.

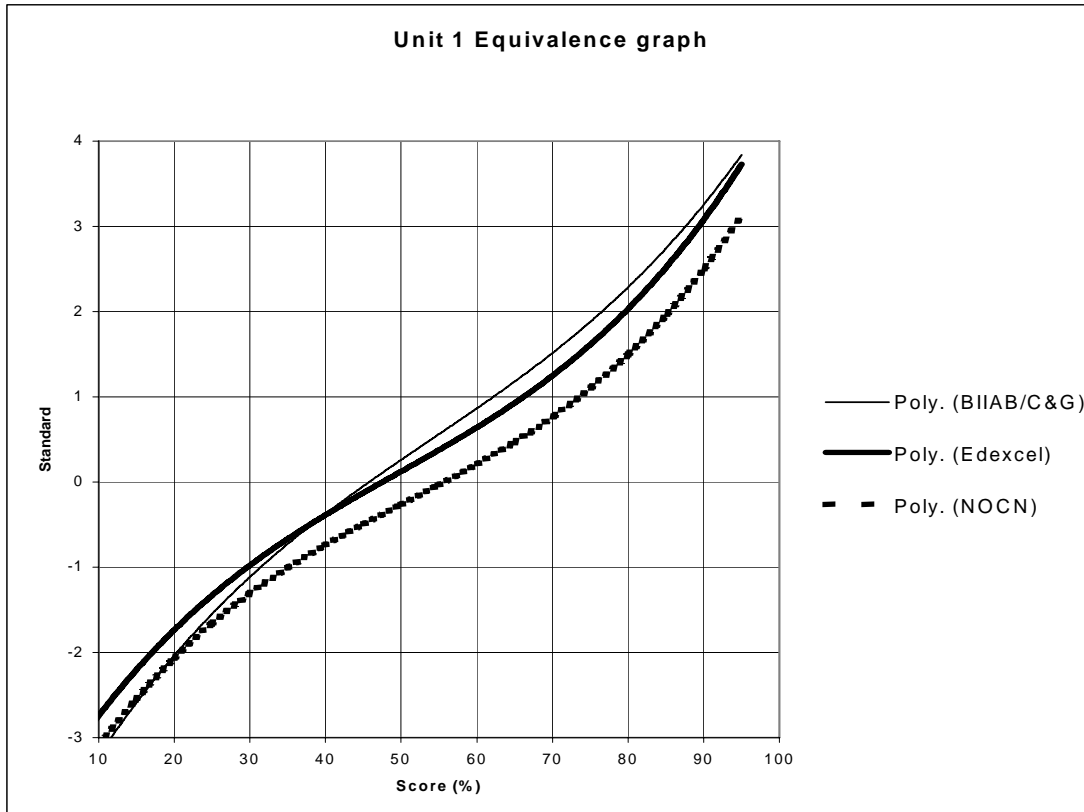
Instead, a procedure was adopted for comparing the Unit 1 tests that combines elements of the standard setting procedures commonly used for multiple-choice tests (such as 'Angoff' or 'bookmarking' methods) with the paired comparison methodology generally used in comparability studies of GCSE, GCE and GNVQ standards.

Four judges, all with general experience of the school/FE/HE system, but with no specific knowledge of these qualifications, were given five sets of questions from the three tests. Each set contained four questions randomly selected from each test and the judges were asked to sort the twelve questions into a rank order of relative difficulty, using their judgement of how difficult each question would be if they had just taken a relevant training course. Then, two new sets of 10 questions were assembled, each containing two questions from each of the original sets, and the judges were again asked to rank the questions by difficulty. This allowed all of the data to be merged into a single data set. Because the judges report *relative* rather than *absolute* difficulty, the method is not affected by any variation in how many questions they think the candidates 'ought' to get right – that is, their data do not judge the standards but just the relative standards of the tests. Experience with similar methods in investigating general qualifications suggests that the use of non-experts as judges is not inappropriate and may be, in some respects, better than using teachers or examiners.

A general problem with this procedure, however, is that the judges cannot accurately imagine how difficult the questions will be for real candidates. While the comparison methodology 'cancels out' all errors in the standard the judges expect, they may think some kinds of question *relatively* more difficult than the real candidates actually find them. If the tests differ in their use of certain types of question this may lead to inaccuracies in estimating their actual relative standards. This issue of the difficulty of particular kinds of questions will be addressed several times in the report, both from the perspective of empirical difficulty and of validity. 'Empirical difficulty' is defined only in terms of how many people get the question right, irrespective of what makes it difficult to get it right.

Comparative data of this kind are analysed in line with Thurstone's *Law of Comparative Judgement*, using Rasch measurement theory. This produces a scale of difficulty on which every item used in the exercise is located. Selecting just those from each test allows a graph to be drawn relating 'score' to 'standard'. Since the exercise used 20 questions from each test, this graph is re-scaled into a percentage score with which to compare the tests. The result is shown below.

Figure 5: Score versus Standard for each test



To report the apparent standards of the tests the NOCN and BIIAB/C&G pass marks will be compared to the Edexcel pass mark; this should **not** be taken to imply that the Edexcel test standard is in any way 'correct'.

The table below shows the procedure. The Edexcel pass mark (pm) is first converted (via percentages) to a 'standard' from the graph by reading vertically from 70 per cent, and the corresponding percentages from the other tests are then read from the graph by tracing horizontally to the other lines and then reading down to the percentages. These percentages are converted to give raw score marks equivalent in standard to the Edexcel pass mark.

Edexcel pm = 70%, which means that the 'Edexcel standard' = 1.247

Table 32: Differences in difficulty

Test	Equivalent %	Equivalent score	Actual pm	Difference
NOCN	76.9	26.9	27	+0.1
BIIAB/C&G	66.0	26.4	28	+1.6

Interpretation

- There is very little apparent difference in difficulty standard among these three tests.
- It is difficult to be sure of the accuracy of this analysis, but experience suggests that a difference of more than about 3 marks may be considered real.

9. Demands, difficulty and validity

9.1 Cognitive demands

In assessment, the level of cognitive demand refers to the nature of the cognitive processes required of candidates in the process of answering the questions. It does **not** refer to the amount or nature of study required in preparation for the examination, nor to the amount of effort or time the examination takes.

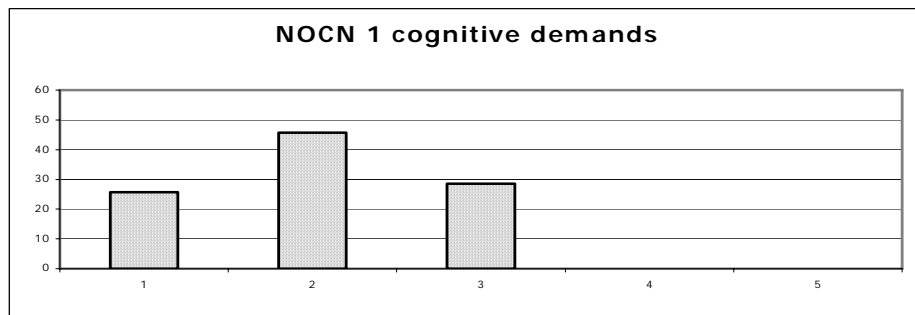
For this exercise a five-level scale was created, with the levels defined as follows:

- 1 Simple fact recall OR simple logic or complex recall made easy by options
- 2 Complex recall, including definitions
- 3 Show understanding of a meaning. Simple options or complex recall made difficult by options
- 4 Show understanding of a meaning: complex options
- 5 Apply reasoning with knowledge or show understanding made difficult by options OR judge relative value of options.

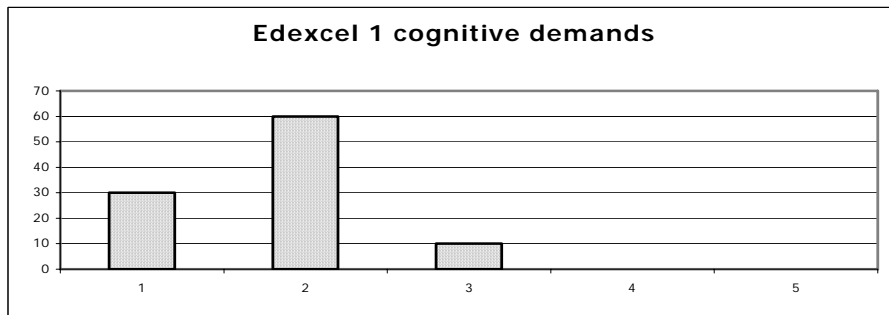
The ratings are shown below.

Unit 1

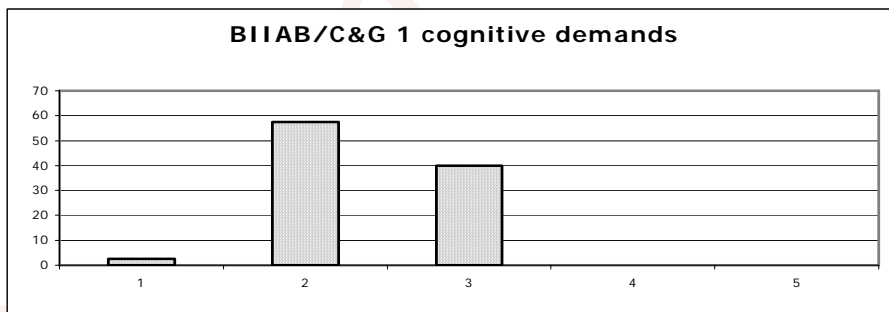
Figure 6: Levels of cognitive demand – Unit 1



Mean demand rating = 2.03



Mean demand rating = 1.80



Mean demand rating = 2.38

The variation in the mean ratings was tested for statistical significance.

Table 33: ANOVA Table for demand ratings – Unit 1

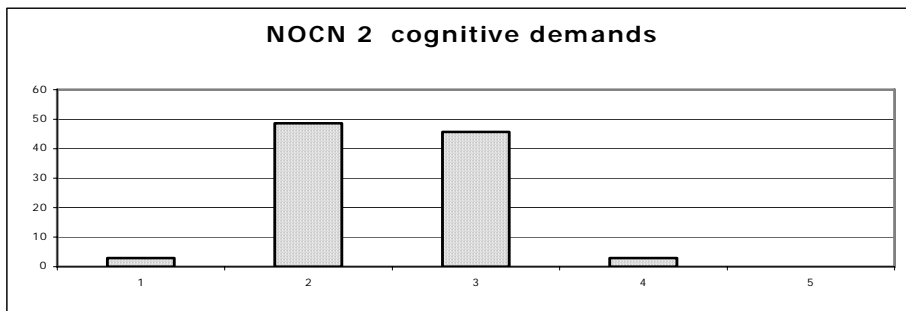
	DF	SS	MS	F-Value	P-Value
TEST	2	5.901	2.951	7.314	.0011
Residual	102	41.146	.403		

In conclusion, the difference between the tests is highly significant: the probability of it being a chance effect is about 1 in 1000.

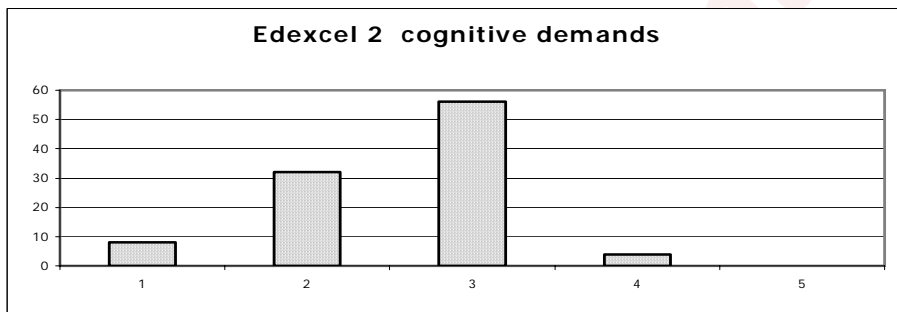
The BIIAB/C&G test asks more questions, and the Edexcel test fewer questions, that require candidates to show that they understand – rather than merely remember – what they have learned.

Unit 2

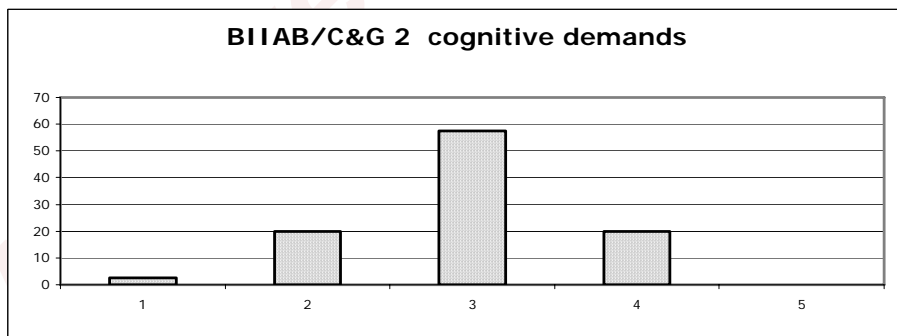
Figure 7: Levels of cognitive demand – Unit 2



Mean demand rating = 2.49



Mean demand rating = 2.56



Mean demand rating = 2.95

The variation in the mean ratings was tested for statistical significance.

Table 34: ANOVA Table for demand ratings – Unit 2

	DF	SS	MS	F-Value	P-Value
DS2 AB	2	4.587	2.294	4.966	.0088
Residual	97	44.803	.462		

In conclusion, the difference between the tests is again highly significant: the probability of it being a chance effect is less than 1 in 100.

Again it is the BIIAB/C&G test that requires more of the higher cognitive skills. Items rated 3 or above (ie understanding and application of understanding) account for 60 per cent of all the items in the BIIAB/C&G test, 40 per cent of the items in the NOCN test, and 35 per cent of those in the Edexcel test.

Interpretation

The questions in the tests vary considerably in regards to the cognitive processes being tested, with the BIIAB/C&G test being most demanding and the Edexcel test being least demanding.

Before concluding that the BIIAB/C&G test is more *difficult* than the others, however, the difference in format of the Unit 2 tests needs to be considered. The BIIAB/C&G test is based on a video recording of four simulated incidents at a club. In each, candidates read questions then watch the recording, before answering the questions and then watching the recording again. The result is a test that is high in (simulated) authenticity, and the candidates are asked to judge the behaviour of the door supervisors they see. It is no surprise that the questions earn higher ratings for 'cognitive category', since it is easy to ask them to evaluate the 'performances' they watch and to imagine themselves in a similar situation.

In contrast, the Edexcel True/False items are rated low in cognitive demand, because they mostly seem to test simple knowledge. But the format artificially increases the complexity of the test task by greatly increasing the amount of logical reasoning that candidates feel is needed. For example, it is superficially simple for a candidate to disagree with:

Opinion evidence is that of a qualified person

a. True b. False

but the format leads some candidates, especially the more anxious ones, to doubt the simple answer and wonder about distinctions between 'qualified' and 'expert', or 'opinion' and 'expert', and to wonder what the point of the question is. True/False encourages the more suspicious candidates to look for tricks in the question, and so is never as simple as it seems.

It is fairer, then, to think of the BIIAB/C&G test as more authentic, rather than more demanding, than the others. It is clearly good, in principle, to use video recordings as the basis of a test of skills such as those expected of a door supervisor, because more sophisticated thinking processes can be assessed without making the testing process overly complicated. Edexcel also tries to base part of its Unit 2 test on the concept of scenarios, but presents the scenarios in writing. This seems much less satisfactory because of the considerable amount of extra reading that is required to present a story and because it is more difficult to understand a written scenario than a video. Specific words in a written narrative cue specific responses in a way that a video film will not do. In

contrast to these two approaches, NOCN avoids scenarios almost completely; a few of the items do set a context for a single question, but this happens as much in Unit 1 as in Unit 2.

It is worth considering together the findings on demand from this section with those on difficulty from Section 8. Overall, the empirical exercise found little difference between the tests in Unit 1 in terms of how difficult they would be to pass. The Edexcel test is more at risk, in that candidates are more likely to pass through luck. The NOCN test is the most dependable in that respect. Also, the Edexcel Unit 1 test relies most heavily on what might be considered simply the ability to remember facts, and on True/False items. All three tests try to evaluate understanding and skills more in Unit 2. NOCN does this with standard four-option multiple-choice items. Edexcel uses more True/False items, but adds some written scenarios with four four-option items on each. While BIIAB/C&G uses video recordings with 10 three-option items on each. The results are:

NOCN: the tests are probably highly reliable, and certainly are not susceptible to random guessing. They test quite well what discrete multiple-choice items can easily test. The test, especially in Unit 2, is largely concerned with knowledge and understanding of the many models (eg POP, SAFER, Lead & Support, Open PALMS) and jargon (eg 'dynamic risk assessment', 'high risk') that are taught in the course, but is not obviously a convincing test of the ability to use them appropriately.

Edexcel: the use of True/False two-option items and the preponderance of low cognitive level items, especially in Unit 1, makes the test rather susceptible to chance and of questionable reliability. It is rather easy to pass the tests without much knowledge from the course if the candidate can cope with the heavy reading load.

BIIAB/C&G: the Unit 1 test is quite similar to the NOCN test, but puts more emphasis on understanding and its application; there were fewer invalid clues in its items than in any other test. In Unit 2 the video provides a test that seems more valid to the candidates, and which allows the examiners to ask sophisticated questions without overloading the candidates. On the other hand, the role that chance can play in it is worrying. Rather than trying to dream up fourth options for the items, BIIAB/C&G should consider where the pass mark should be set. A careful study of pass rates and who passes and fails in the two units will ensure that this is set at an appropriate level.

9.2 Test format demand

Assessment procedures differ across the three qualifications. In this section these procedures are compared. The key issues here are test length, reading demand and scenario-based assessment. These issues are inter-related. All other things being equal, candidates would doubtless prefer a short test to a long one, but they might also expect more time if the test requires a lot of reading.

Basing items on scenarios is likely to change the amount of reading required and may affect the complexity of the test procedure.

The details of reading demand and time for each test and subtest are as follows:

Table 35: Overall tests

Examination	Items	Words	Words per item	mins	Rate (items per min)	Rate (words per min)
NOCN	70	2,970	42.43	120	0.58	24.75
Edexcel	55	2,099	38.16	120	0.46	17.49
BIIAB/C&G	80	3,080	38.50	120	0.67	25.67

The word count here is the total number of words in the stem plus all the options in each question. Any rubrics not part of the questions are not included.

The Edexcel tests are, overall, considerably less demanding than the other two, both in terms of the number of items to be answered in two hours and in terms of the overall amount and rate of reading required. This may be appropriate given the nature of the candidates for this qualification. The nature of the reading demands becomes clearer if we look at these statistics for units and subtests.

Table 36: NOCN unit tests

NOCN	Items	Words	Words per item	mins	Rate (items per min)	Rate (words per min)
1: Multiple-choice	35	1,390	39.71	60	0.58	23.17
2: Multiple-choice	35	1,580	45.14	60	0.58	26.33

The Unit 2 test has slightly more words than the Unit 1 test, probably because the options in Unit 1 are more factual, and so more easily expressed in a few words. The demand of 0.58 items per minute is, of course, the same in both units.

Table 37: Edexcel unit tests

Edexcel	Items	Words	Words per item	mins	Rate (items per min)	Rate (words per min)
Edexcel 1	30	811	27.03	60	0.50	13.52
Edexcel 2	25	1,288	51.52	60	0.42	21.47

These are relatively short tests, at 30 and 25 items to be completed in one hour, compared to the 35 and 40 items to be dealt with in the tests given by other two awarding bodies within the same amount of time. The Unit 1 test is also very light in its demand on reading speed, with about half as many words per minute as the NOCN tests. The Unit 2 test is lower in overall reading demand than the NOCN tests, but the items themselves contain more words. These figures are clarified when we look in more detail at the item types used.

Table 38: Edexcel Unit 1 sub-tests

Edexcel 1	Items	Words	Words per item
True/False	10	142	14.20
Multiple-choice	20	669	33.45

Edexcel makes use of the True/False format. Whatever else may be said of it, it is very economical in reading terms, with an average here of just 14.2 words per item. The multiple-choice items, at 33.45 words per item are still shorter than the average NOCN item.

Table 39: Edexcel Unit 2 sub-tests

Edexcel 2	Items	Words	Words per item
True/False	4	49	12.25
Multiple-choice	5	195	39.00
Scenarios	16	1,044	65.25

The figures here show the problem of the 'verbal scenario'. While True/False items are again used, there are just four of them, together with five multiple-choice items, which are again shorter than NOCN's Unit 2 items. If the aim is to reduce the reading load, however, this is defeated in the 'scenario' section in which four scenarios are presented verbally, using a total of 142 words in the initial presentations. In addition, each scenario is developed in the stems of the following items, leading to an average of 65.25 words per item, easily the highest reading demand of any sub-test.

Table 40: BIIAB/C&G unit tests

BIIAB/C&G	Items	Words	Words per item	mins	Rate (items per min)	Rate (words per min)
1: Multiple-choice	40	1,656	41.40	60	0.67	27.60
2: Scenarios	40	1,424	35.60	60	0.67	23.73

The BIIAB/C&G Unit 1 items are typical of all multiple-choice tests in using about 40 words per item; the Unit 2 items are shorter only because they use three instead of four options. The tests are slightly longer than NOCN's for the same time allowance.

Discussion

The biggest difference between the tests, from the candidate's point of view, is the use of video by BIIAB/C&G, verbal narrative scenarios by Edexcel, and no scenarios at all by NOCN.

The Edexcel tests are paradoxical in that, for the most part, they aim to minimise the reading demands on candidates (this aim was confirmed in interview) yet they present scenarios verbally, which greatly increases this demand. The overall low reading demand comes largely from the use of the True/False format but, as discussed earlier in the section on cognitive demand, this substitutes a demand on reasoning strategies for the reduction in reading demand. It is not clear, therefore, that the Edexcel tests succeed in reducing the level of 'construct-irrelevant' demands on candidates. (This term is meant to include all demands on candidates that are not closely

identifiable with the knowledge and skills that the training is meant to develop and the qualification to ensure, and in this case it includes reading and logical reasoning abilities.)

In contrast, BIIAB/C&G minimises the construct-irrelevant demand of narrative reading by its use of video in presenting the scenarios that support the items in Unit 2, replacing the need to read with the need to watch and listen. It is easy for all candidates to follow the narrative in video mode. There may be some concern that non-native speakers could find it more difficult to understand the speech in the recordings without seeing it in print, but this might also be viewed as a valid difficulty since it reflects actual work conditions. The videos are shown three times, and the procedure is clearly signalled to candidates, which should ensure that they all have adequate opportunity to consider every question in the narrative context. The use of just three options seems to make the items rather easy, and in several items it seemed that some options could be discounted for reasons not related to the relevant skills. BIIAB/C&G can check the overall difficulty of this Unit by monitoring the pass rate in comparison to Unit 1, and should do so to ensure that the pass mark is appropriate.

NOCN makes no attempt to base its Unit 2 test on scenarios. Instead, it sets standard multiple-choice items that test knowledge of models for appropriate behaviour. It must be questioned whether the multiple-choice format it uses is capable of assessing candidates' *application* of their learning or of the skills of door supervision, rather than merely *knowledge* of what a door supervisor ought to be able to do.

10. Telephone interview findings

Appropriate people in each of the awarding bodies were interviewed by telephone. These were used to check various issues that had arisen during the analyses, and to explore the procedures they each used to write and review questions, to construct the tests, to determine pass marks and to monitor question quality. Questions were also asked about development plans.

10.1 Question writing

NOCN: There are two groups of item writers operating in the broad areas of general security and the security of cash and valuables. For each paper in this qualification a single writer is commissioned to write the items. In normal circumstances, a sample of these items is then considered by a review team consisting of the four writers in that group and four others with expertise in assessment, quality assurance and the subject matter. Last year, because the qualification was new, all of the items were reviewed.

The item writers received training from NOCN.

Edexcel: There are two item writers, who have been trained by Edexcel. For each Unit, one of these acts as 'lead writer' and the other as 'editor'. In addition, all items are reviewed by a chief examiner. A pilot examination was run before the qualification was launched, to explore suitable item types; these items were not retained, but new items were written similar to the successful pilot items.

BIIAB/ C&G: Items are written by a team with teaching and legal experience, and reviewed by a separate team of two experts and two BIIAB/C&G staff against assessment outcomes and other quality criteria. They are then piloted, and statistics and comments from centres are considered before they are added to the bank.

Scenarios for the Unit 2 test were selected from the SIA list of potential scenarios and scripts written for the actors. Questions were written to accompany the script, and revised after recording. The decision to use three options (rather than four or two) was arrived at after much discussion with training deliverers and centres.

All three awarding bodies have adequate systems in place for item writing, though the high number of language errors in the Edexcel tests (Section 7 '*Question difficulty and quality*') suggests that their scrutiny process needs reviewing, and the high 'plausibility' scores for all six tests (Section 7) suggest that all of the teams would benefit from more training in item writing.

10.2 Test production

NOCN: At present there are seven papers in operation; each one stays live for a period of about six months, until it is considered to have been exposed enough and is withdrawn. Some items are used in more than one test, and the total number of items that may be considered to be in a 'bank' is around 250 for each unit. The aim is to increase the size of the bank so that no test will remain live for six months and so that there will be no need for overlapping items. Skills for Security (SFS) constructs the tests and has responsibility for ensuring content validity. For Unit 1, the test specification requires three or four items in each of 11 'modules'. It is the narrowness of subject matter of some of these modules that makes it difficult to write many items, leading to re-use of items in different simultaneous tests.

Edexcel: After review, the items are banked according to learning outcomes, or modules. There are four of these for each unit. Item banking software then selects a minimum of five items from each to produce a paper for each unit and these are reviewed by a chief examiner as the two components of an examination. The chief examiner checks coverage and looks for overlap or other conflicts between items. Items are re-used, but not in tests that are simultaneously live; five tests are currently live for each unit.

The banks hold about 300 items for each test. The whole bank is reviewed annually; specific changes to legislation or regulation also prompt reviews.

BIIAB/ C&G: Unit 1 – Items go into a bank after review. There are currently about 300 items being used in live tests. Many items have recently been retired because of changes in licensing and security law and 86 new items have been added in recent months. A draft paper is generated by computer to fit the specification and is then checked for consistency and overlaps.

Unit 2 – There are 11 video tapes at present, each containing four recordings of incidents, with at least one from each of the three SIA ‘types’ of scenario. Because test construction consists only of choosing four sub-tests for each test, from the 44 available, the task is not automated.

The SIA specification states that:

The questions for examination papers will be selected at random from a question bank held by each awarding body. (SIA: Section 2)

The reason for specifying random selection is, presumably, that randomly equivalent tests can be assumed to be more or less equally difficult, but if taken literally, the randomness requirement would seriously weaken the content validity of the tests. As this qualification is fairly new, items go into tests quite soon after being written, and the three awarding bodies could be described as being in the process of setting up their item banks. Edexcel and BIIAB/C&G use a system that involves some randomness in the selection of items within sub-tests for Unit 1, and NOCN will also be able to do this when they have enough items. Random selection *within* subtests, rather than for the whole ‘paper’ should be adequate to maintain standards as well as content validity.

The randomness demand is inappropriate for the scenario-based testing implied by the specification and used by Edexcel and BIIAB/C&G for Unit 2, where whole sub-tests of four or 10 items must be selected together. Especially for BIIAB/C&G, the need to avoid overlap between sub-tests means that human control is essential. It would be unfortunate if regulation imposed a testing method that is not obviously able to assess ‘*appropriate application* of the knowledge, skills and understanding’ (SIA: Section 2) and prevented the awarding bodies from developing more imaginative – and popular – procedures.

10.3 Pass marks and item/test statistics

NOCN: SFS set the pass marks at 27 out of 35 for each of the two tests. Centres are encouraged to report on any difficulties they see in items. This is the main routine procedure for monitoring item performance.

Edexcel: Pass marks were set at approximately 70 per cent (21 out of 30 and 17 out of 25) for all tests. The main post-test analysis is a monitoring of pass rates within centres (partly as a way of spotting possible malpractice) and across tests; as a result of this the pass mark on any one test could be changed, though this has not yet happened.

BIIAB/ C&G: The pass mark is constant across tests within each unit, and was agreed by the moderating team following discussions with a wide group of interested parties, including those delivering training, content experts, test users and magistrates. Feedback is invited from centres.

Simple item analysis is run approximately monthly. A report for each item shows how responses are distributed across the key and distractors, and is scanned for evidence of problems. Pass rates are being monitored for each unit test, and pass marks may be adjusted if the rates prove to be out of step.

Careful monitoring of results, especially of pass rates on different tests, is the only method available to the awarding bodies for ensuring comparability between tests in Unit 1. This will change when the item banks are large enough to make random equivalence effective.

For Unit 2, pass rate monitoring will always be the most appropriate method for establishing the pass marks. Should the pass rate be the same in both units, for the same candidates? In the absence of any strong argument to the contrary this would be the default assumption, but it is not explicitly stated in the specification that the two units should be of equal weight in determining results. If Unit 1 tests are randomly equivalent and Unit 2 tests prove to have the same pass rates as Unit 1, then the Unit 2 tests will also be randomly equivalent. There will then be no need for random item selection to ensure equivalence in Unit 2.

10.4 Plans for future development

NOCN: The tests are being developed for administration on-line, partly in response to the wishes of the largest single customer for the qualification, and partly in order to improve test security. Two pilots of assessment on-line have been carried out.

Ultimately, evolution of the system is planned so that it will move to assessment on demand rather than at set administration times.

Edexcel: Of the three awarding bodies, Edexcel is closest to running a genuine item banking system, with random item selection to meet a specification and re-use of items. The missing element is item analysis, which would provide better checks on item quality and a better guarantee that the tests maintain the same standard. Item analysis is planned.

An on-screen version of these tests may soon be available to be taken on-line through Pearson's network of VUE centres. Web-based testing has been considered, but is unlikely to be appropriate for this qualification.

BIIAB/ C&G: Unit 1 – The current system is a hybrid, perhaps, best described as a test bank in the process of evolving into an item bank. Plans centre around its development into a full item bank system supporting e-assessment and an internal pilot has been carried out. The awarding body is alert to the potential of a bank for maintaining flexibility, as legislation and practice force changes in the tests.

Unit 2 – The video-based test format is popular with centres and is likely to be retained. Additional scenarios will be developed to increase the range, but the assessment will remain as a bank of tests rather than a bank of items.

10.5 General comment

All of the awarding bodies are developing their systems to exploit the potential of fully functional item banking to raise assessment quality and to ensure constant pass standards. To make the most of this, item analysis based on live data will need to be an integral part of the system, so that poor quality items can be screened out, and candidates are protected from the chance of meeting an unusually difficult, or easy, combination of items. A bank supporting many simultaneous test versions, or an on-screen system for delivering randomly equivalent tests to different candidates, is probably the best way to guarantee the security of test papers or on-screen tests.

11. Overall strengths and weaknesses

11.1 Strengths

The following may be considered as strengths in some of the tests analysed:

Content validity

All three of the awarding bodies ensure that their tests closely match a content specification, although there are significant differences in how they do this.

Standards

An empirical study showed no evidence of a difference in pass standard between the three Unit 1 tests. No such study of Unit 2 was possible.

Item quality assurance

Training in writing skills seems adequate in all the awarding bodies. In one case it is not so clear that the procedures for catching faults by review would generally be adequate.

Item banking

All of the awarding bodies are moving towards fully operational item banking systems. Although they differ, all of the systems at present are reasonably good and, as they develop, will be better able to ensure that standards are maintained. Scenario-based assessment, where used, does complicate the operation of item banking. All awarding bodies are also considering on-screen testing.

11.2 Weaknesses

Some weaknesses were also noted in the construction of the tests. Attention to these would improve the quality of assessment.

Reading difficulty

The reading demands of the tests sometimes seemed too high for the expected candidates. Sometimes, this resulted from using the formal language of legislation, which is difficult for many Level 2 candidates. In other cases, unintended syntactic complexity or unnatural phrasing cause an invalid source of difficulty.

Language errors

There were many errors in the spelling, punctuation or syntax of items in the tests from one awarding body, but few from the others.

11.3 Strengths and Weaknesses

In some respects the tests showed both good and bad features.

Scenario-based assessment

The Security Industry Authority (SIA) lays down an outline specification that implies that Unit 2 should be based on learning appropriate responses to various 'scenarios'. One of the awarding bodies uses video recordings of acted scenarios to assess this unit, one uses a few scenarios described in writing, the third does not use any scenarios.

Item quality

A person with adequate reading skills could expect to get around 50 per cent of items correct without specific knowledge. Two awarding bodies use two- and three-option items, which are easy to guess, with the consequence that a weak candidate has too high a chance of passing through luck.

On the other hand, there were many technically good items in the four-option multiple-choice sections and some imaginative attempts to assess skills in Unit 2.

Cognitive demand

In Unit 1 the tests vary significantly in their balance between questions that evaluate recall of facts and those that evaluate understanding. In only one test do more than one-third of the questions address understanding. There are large differences too in Unit 2, which is explicitly meant to be about understanding and application of appropriate skills.

Test format demand

The tests vary considerably in format demand, both in terms of the number of items to be completed per minute and the number of words to be read per minute, required of candidates.

Item statistics

The awarding bodies vary in their approach to monitoring item performance statistically, and only one does so routinely. To implement a more developed item bank system the others will need to develop more efficient and automated ways of doing this.

12. Conclusions

- The awarding bodies differ in how they interpret the content demands of the test specifications. For Unit 1 it is not clear whether this affects content validity. Empirical investigation shows little evidence of differences in standard between the Unit 1 tests.

- The awarding bodies take very different approaches to the design of the Unit 2 tests. NOCN tests mainly the learning of models and theories for appropriate behaviour. BIIAB/C&G tries to test the application of learning fairly directly, with tests based on videos of simulated incidents. Edexcel also bases some of its test on scenarios, but expresses these verbally in quite long written question rubrics.
- A review should be carried out by, or including, an independent content expert, to consider whether these tests meet the SIA requirement to assess the *application* of skills in Unit 2. It should also consider whether the models and content being taught is sufficiently up to date.
- Several of the faults commonly described in textbooks and training materials for multiple-choice item writers were common in these tests. Procedures for catching these should be tightened.
- A few options in Edexcel multiple-choice items were judged 'not possible' right answers.
- In all of the tests, an intelligent but untaught candidate could expect to score around half marks – a little less in the BIIAB/C&G test and a little more in the NOCN and Edexcel tests.
- The length of options gave invalid clues to the right answer in about 10 per cent of the NOCN and Edexcel items.
- Reading difficulty was judged to be a possible invalid source of difficulty in a significant number of questions, especially if candidates are not native speakers of English. This was more prevalent in NOCN and Edexcel questions than in BIIAB/C&G ones.
- The NOCN tests are more secure than the others in their standard, in that there is little chance that a candidate will pass without an adequate level of knowledge. The BIIAB/C&G Unit 2 test is particularly sensitive to the effect of luck.
- Text highlighting was not used by Edexcel, and rarely by NOCN, to help candidates understand the meaning of items.
- A quarter of the Edexcel items seen contained errors of spelling, punctuation or grammar, and they were inconsistent in the question format used.
- BIIAB/C&G consistently describes door supervisors as male.

- The BIIAB/C&G tests ask more questions that require evidence that candidates understand, and can apply, their knowledge than do the other tests. This is due to BIIAB/C&G's video format in Unit 2, which facilitates the asking of questions that require an evaluation of possible alternatives.
- The Edexcel Unit 1 test is the least demanding in requiring evidence that candidates understand and can apply their knowledge. This is largely because Edexcel sets about one quarter of its questions in a True/False format, which is superficially simple – they are not rated high in terms of cognitive level, but are more difficult than it first appears.

13. Recommendations

For the awarding bodies:

- All awarding bodies should ensure as much and as soon as possible that items for Unit 1 tests are randomly chosen from an adequate bank of items *within 'learning outcomes' or 'modules'*, so that their pass marks benefit from the assurance that 'random equivalence' brings.
- For Unit 2 they should monitor pass rates in comparison with Unit 1, rather than rely on random equivalence.
- A review should be carried out by, or including, an independent content expert, to consider whether these tests meet the SIA requirement to assess the *application* of skills in Unit 2. It should also consider whether the models and content being taught is sufficiently up to date.
- All of the awarding bodies should be encouraged to move quickly towards their declared aims of implementing fully featured item banking systems. This will improve the quality of the items that make up the tests and increase the confidence that certificate users may have in the standard of the qualification.
- Systems for monitoring the quality of items to identify and delete faulty items – both at item review stage (before they are used in tests) and through analysis of real test data – need to be improved.
- NOCN should also consider basing at least part of the Unit 2 test on scenarios, to ensure that it does test the application of learning.

- The Edexcel tests for Unit 1 seem very undemanding in terms of reading, because they use fewer words than the others, but the psychological peculiarities of True/False items make this rather misleading. The decision to present scenarios through extensive written text also makes the Unit 2 test particularly difficult for candidates with poor reading skills. Edexcel should consider other ways of presenting scenarios.

For the SIA:

- The SIA should relax the requirement for random item selection in Unit 2.

Archived Content

Appendix 1: Acronyms

Edexcel

C&G – City & Guilds

NOCN – National Open College Network

BIIAB – BIIAB

SIA – Security Industry Authority

SITO – Security Industry Training Organisation

SFS – Skills For Security

Archived Content

Appendix 2: The SIA test specification*

Topics and objectives

Part 1: Role and Responsibilities of Door Supervisors in the Security Industry Environment

1a Introduction

- Define the role of the door supervisor.
- Identify the qualities of a door supervisor.
- Identify the key players in the leisure and security industries.
- State the relationships with the SIA, the police and local authorities.
- State the main objectives of door supervisors.
- State the requirements for door supervisors under the Private Security Industry Act.

1b Behavioural standards

- State the reasons for having behavioural standards.
- State the standards of behaviour required of door supervisors.
- List the SIA's specific requirements in relation to licensing and enforcement.

1c Civil and criminal law

- Show an understanding of civil and criminal law.
- State the requirements relating to the use of force.
- State what is meant by 'reasonable' and 'necessary force'.
- Identify types of assault.
- List some other criminal offences that door supervisors may come across.
- List the options available when the law is broken.

1d Searching

- State the reasons for searching premises.
- State how to search people and their property.
- State 'conditions of entry' and the importance of obtaining permission to search.
- State the difference between general, random and specific searches.

- Define an 'offensive weapon'.
- State the hazards of conducting a search.
- State the procedures for recording articles seized during searches.
- List the options available to door supervisors if they find items during searches.

1e Arrest

- Define an arrestable offence.
- State the agreed procedures for an arrest.
- State the limitations of a door supervisor's powers of arrest.
- Identify circumstances in which door supervisors may arrest.
- Specify why arrests should only be made as a last resort.
- State the procedures following an arrest.

1f Drugs awareness

- Identify key areas of the Misuse of Drugs Act of 1971 and the Public Entertainment Licence (Drugs Misuse) Act 1997.
- State some of the symptoms of drug abuse.
- List the most common types of illegal drugs.
- List signs of drug dealing.
- State how to deal with customers found in possession of drugs.
- State the procedure for handling seized drugs.
- List health and safety issues in relation to illegal drugs.
- State how to dispose of contaminated waste.

1g Recording incidents and crime scene preservation

- List the types and reasons for records.
- State the reasons for recording incidents.
- Identify incidents that need to be recorded and when to call the police.
- Explain what information a record should contain.
- State the rules for incident book/notebook entries.
- List the different types of evidence.
- State how forensic evidence can be obtained at a crime scene.
- State the basic rules to follow to preserve evidence.

1h Licensing law

- State the law in relation to refusing entry and ejecting customers.
- State police powers in relation to licensed premises.
- State the different types of licences and permissions available for premises.
- State the rights and duties of licensees and door supervisors as their representatives.
- State the law in relation to young persons.
- State law in relation to drunkenness, disorderly conduct, prostitutes and unlawful gaming.

1i Equal opportunities

- State what is meant by equal opportunities, prejudice and stereotyping.
- State the relevance for door supervisors of the Race Relations Act, the Sex Discrimination Act and the Disability Discrimination Act
- State how these Acts may be enforced.

1j Health and safety at work

- Define the main areas of Health and Safety legislation.
- State the purpose of 'duty of care'.
- State the responsibilities of the employer, employees and self-employed.
- Identify typical risks, hazards and spillages.
- Identify safety signs and safety signals.
- State the precautions to be taken in the manual handling of heavy goods.
- State the precautions to be taken against HIV, hepatitis and other infectious diseases.

1k Emergency procedures

- Define the term emergency.
- List the reasons for fire certificates and occupancy figures.
- State the three components of fire.
- Explain the importance of knowing the venue's fire and evacuation procedures.
- List the types of fire extinguishers, and their use.
- State the procedures for a bomb threat.

- Demonstrate an understanding of how to remain alert and vigilant of unusual and suspicious activity.
- State how to act in a first response situation requiring first aid.
- State who to contact in first aid situations.

Part 2: Communication Skills and Conflict Management

2a Introduction to communication skills and conflict management

- State the importance of customer care.
- Explain basic communication skills.
- Demonstrate verbal and non-verbal communication effectively.
- State the need to calm difficult situations and avoid violence.
- Identify different types of behaviour.
- State what risks can occur during violence at work.
- Identify the most common conflict flash-points.
- State the importance of managing aggression.
- Demonstrate basic problem-solving skills.
- State the impact of reflecting and learning from the experience of conflict.

2b Application of communication skills and conflict management

Refusing entry to a customer on the grounds of:

- the venue being full to capacity
- being under the influence of drink and/or drugs
- being underage
- not being suitably dressed
- not being able to pay the entrance fee
- refusal to be searched
- being found in possession of weapons or drugs
- being banned or under an exclusion order.

Ejecting a customer from the venue due to:

- breaches of criminal law (eg theft, damage, assaults, drugs)
- breaches of licensing law (eg being drunk, violent, quarrelsome)
- breaches of house rules (eg dancing on tables, taking bottles onto the dance floor).

Incidents inside the venue, such as:

- advising/reprimanding for behaviour (breaches of house rules)
- first aid situation
- undertaking an arrest of a customer for an arrestable offence
- failing to adhere to drinking-up times
- domestic disputes
- other disputes (eg customer vs bar-staff, complaints about service)
- arguments/fights.

2c Scenario situations

Refusal scenarios

- Refuse entry to customers when the venue is full, explaining why no more people can come in, explaining that they may be allowed in if and when other customers leave.
- Refuse entry to someone who is obviously under the influence of drink or drugs. Some customers may complain, but go away; others will argue and may become aggressive.
- Refuse entry to someone who appears to be under the age of 18 (or 21), and cannot provide appropriate ID.
- Refuse entry to someone who breaches the venue's dress code, ie someone wearing trainers where this is not allowed.
- Refuse entry to someone who wants to come in free of charge, who either cannot pay the admission charge or claims to be a friend of someone who works there.
- Refuse entry to someone who refuses to be searched as part of the entry conditions.
- Refuse entry to someone found in possession of either an offensive weapon or drugs.
- Refuse entry to someone who is banned from entering the premises because of previous behaviour, or who is under a court exclusion order not to enter licensed premises or, who is on a 'pub watch' ban.
- Refuse entry to someone who behaves aggressively at the point of entry and is therefore not suitable to be allowed in.

Ejection scenarios

- Eject a customer for being suspected of theft, criminal damage, assault or drugs inside the venue (where no Police action is required). Some customers will leave when asked to do so; others will argue and/or may become aggressive.

- Eject a customer who breaches licensing laws by becoming very drunk or argumentative or aggressive inside the venue. Some customers will leave when asked; others will argue and/or become aggressive.
- Eject a customer for breaching a house rule such as repeatedly dancing on tables or carrying bottles/glasses on the dance floor.

Incident scenarios

- Advise a customer regarding unacceptable behaviour inside the venue. Try to stop the behaviour, warning the customer about further action if the behaviour persists.
- Deal with a first aid incident where other drunken customers try to take over.
- Deal with a domestic dispute which turns into a noisy incident inside the venue.
- Deal with other disputes inside the venue, such as a customer arguing with bar-staff over incorrect change given, or a complaint about poor customer service.
- Deal with various aggressive arguments between customers to try to prevent them from turning physical.
- Deal with customers (usually drunk) who refuse to leave the premises at closing time.
- Arrest a customer as a last resort. Some customers will comply with the arrest; others will become argumentative or aggressive.

* This specification is dated January 2006