

## Acceptance and Usage of e-Assessment for UK Awarding Bodies





## Contents

Research Stakeholders	3
Executive Summary	4
Introduction	5
Project Theme	6
Purpose and Scope	7
Methodology	8
Qualitative Findings	10
Quantitative Results	11
Overview	26
Conclusions	28
Appendix	29
References and Further Reading	37

## Research Stakeholders

Acritas was pleased to be re-commissioned for this continuing, comprehensive study. By maintaining the same methodology as in 2005, detailed feedback and comparable findings have been delivered, providing both an up-to-date insight into current views of e-Assessment and valuable time-series data highlighting shifts in perception. We were delighted with the very high rate of participation by respondents who have, once again, shown their support of this important project by giving up their time and offering thoughtful and open contributions. It is due to the commitment and enthusiasm of all parties involved that this successful project continues to be an industry benchmark.

*Victoria Riccalton - Acritas Research*

This research is a cornerstone of Thomson Prometric's significant investment in the markets in which we operate. The study shows the changes that have taken place in key areas, but also highlights new emergent areas of interest such as the use of psychometrics and multiple choice in higher learning levels as Awarding Bodies continue to adopt e-Assessment.

It is clear that many of the barriers to adoption that were called out in last year's research have been overcome, resulting in higher levels of acceptance and usage. Comparisons with the results from the previous study make for fascinating reading as the pace of market change increases. We will continue to invest, contribute and listen to market issues, based on these findings, so that we can offer and tailor solutions that support the goals and objectives of our prospects and clients.

Thomson Prometric is proud of our association with this research and we thank all those assessment professionals who gave their precious time to make a valuable contribution.

*Geoff Chapman - Thomson Prometric*



## Executive Summary

This research provides an exploration of the UK e-Assessment market, in relation to the UK Awarding Bodies, comparing findings with those of twelve months ago. It also elucidates on the key areas that have emerged since the first research was conducted. This in turn has provided an insight into the remaining drivers and barriers to the adoption of e-Assessment, but also the widespread acceptance and adoption in the UK.

With 81% of all recognised Awarding Bodies being interviewed, this study is verging on an Awarding Body e-Assessment census based on sound research principles which will lead to continuing e-Assessment development.

It is apparent that the level of e-Assessment industry knowledge and uptake of programs within UK Awarding Bodies is at a much more advanced position compared to the previous research findings. The pace of market change has clearly quickened. It is possible to state that these findings will allow Awarding Bodies to revisit their thoughts on e-Assessment, altering the pace of market maturity in the short to medium term.

Questions related to topics such as psychometrics, use of multiple choice questions for higher levels of learning and e-Assessment location preference, have provided responses which give a sign-post for the key emergent needs of this market.

*There is a significant increase in UK Awarding Body e-Assessment uptake and knowledge*

*Emergent areas such as psychometrics and extended use of multiple choice questions are recognised*



*The regulator has called out how e-Assessment can help the learner and what issues Awarding Bodies might face*

## Introduction

The UK assessment industry continues to experience changes that resonate with learners, educators and the media at large.

The first research study used the QCA's 'Blueprint for e-Assessment' as a clarion call for UK Awarding Bodies to review their e-Assessment provision. Dr. Ken Boston, QCA's Chief Executive, talked about the need for QCA to "*embrace a technological future for..assessment*" and how QCA want to "*vigorously drive forward towards early achievement of this vision of assessment.*"

Dr. Boston specified the following challenges by 2009:

- *All new qualifications will include an option for on-screen assessment.*
- *All Awarding Bodies should be set up to accept and assess e-portfolios.*
- *Qualifications specifically designed for electronic delivery and e-Assessment, should be developed, accredited and live.*

The concluding 'line in the sand' from Dr. Boston was that "*by 2009, e-Assessment should certainly be normal, if not the norm, for thousands of students each year.*"

Dr Boston's 'Strategy, Technology and Assessment' speech at the Victorian Curriculum and Assessment Authority of Australia Roundtable, fleshed out more detail on the e-Assessment vista and how the regulator considered how technology would benefit the learner and that e-Assessment was 'worth the candle'.

With the QCA's original deadline less than three years away, and the regulator calling out its benefits and the market landscape, many Awarding Bodies are now at a planning and change management crossroads regarding their e-Assessment adoption, or fine-tuning their current programme. The research will document these experiences and also the emergent issues that have been discovered since the last study.



**Both newly emergent  
and existing  
e-Assessment issues  
would be documented  
and inform the  
quantitative research**

## Project Theme

The findings of the 'Drivers & Barriers' research gave forth to discussion regarding the measurement and rate of market change, and key facets of e-Assessment that have emerged in the last twelve months.

Using the previous study as a benchmark and noting the changes in the regulatory environment and further exploration of e-Assessment issues by the QCA, it was decided to consider the acceptance and usage of e-Assessment.

Ken Boston's 'Strategy, Technology and Assessment' speech, affirmed that the regulator was committed to e-Assessment, understood the market landscape and how it was changing.

Thomson Prometric decided to commission the research that would uncover the reasons for adoption, how deeply it was ingrained into the UK Awarding Bodies and flag emergent issues. Whilst acknowledging the Blueprint for e-Assessment from 2004, ten qualitative interviews were undertaken with key, major players that would bring forth common and newly emergent themes that would help to frame the quantitative research.

The specific areas of research would include:

- *Detail the extent of feelings towards the Blueprint and the changes therein.*
- *Detail the Awarding Bodies perceptions of e-Assessment.*
- *Highlight the spirit of co-operation and/or tensions that exist and/or are generated when Awarding Bodies and technology providers work in partnership.*
- *The risks associated in introducing and using e-Assessment.*
- *The emergent issues that require solutions so that Awarding Bodies are in a position to provide e-Assessment solutions that are valid, reliable and fit for purpose.*
- *Documenting Awarding Body feelings towards QCA's ongoing efforts to help and/or facilitate e-Assessment adoption.*
- *The factors that might accelerate movement towards e-Assessment adoption.*



## Purpose and Scope

The study would once again consult the QCA recognised UK Awarding Bodies and other key stakeholders as to their level of acceptance and usage of e-Assessment. Within this market, there are a handful of UK organisations, outside of the UK Awarding Body field, who have great e-Assessment experience and use 'mature systems.' Similar to the previous study, it was important to capture the input of these organisations at the qualitative stage, so that the quantitative phase could be as fully informed as possible.

The idea of a 'census' of the 115 UK Awarding Bodies recognised by QCA was retained with the contacts being those who have a specific responsibility for their organisation's exam or qualification system.

*To ensure comparability  
with the previous  
research, the study  
focuses on those  
responsible for the exam  
or qualification system  
within UK Awarding  
Bodies recognised by QCA*



*An identical methodology was used to maximise comparability and a similar sample size was secured*

## Methodology

Utilising the same methodology as the previous study, in-depth face to face interviews with Awarding Bodies and partner organisations were conducted to understand the major issues within the e-Assessment arena and promote discussion. The qualitative results would then inform the following quantitative stage. Scoping workshops, a project plan and a topic guide were then prepared for the qualitative interviews.

### Statistical Confidence Levels

In order to maintain the statistical validity established in the first study, great efforts were made to secure a similar sample size for the quantitative part of the research. As this has been secured, the confidence levels are the same as the previous study i.e. the findings are statistically within 5% of what the entire 115 Awarding Body population would say. For example, if 70% answered yes to a question, between 65% and 75% of the 115 Awarding Bodies would give the same response.



*Key topics were generated by the qualitative interviews which informed the quantitative questionnaire*

#### **Qualitative Interviews**

Ten qualitative, face-to-face interviews were performed (at the respondent's office) with Awarding Bodies, a regulator/Awarding Body and UK governmental bodies. The interviews call out stakeholder issues in detail and aids understanding of how the e-Assessment marketplace is changing.

The topic guide was strictly followed to ensure all important areas were covered, whilst allowing enough flexibility to probe interesting emergent issues and ensure optimum contribution from each respondent.

Where respondents authorised attributable comments, individual interview reports were prepared and transcribed. Upon completion of the interviews, the output was a written report of findings and a short overview for discussion.

#### **Quantitative Interviews**

Informed by the qualitative research findings, a structured questionnaire was then designed, employing both open and closed questioning techniques in order to combine the benefits of detailed verbatim feedback, together with measurable data.



*93 respondents out of 115 Awarding Bodies (81%) took part in the study*

## Qualitative Findings

From the interviews, seven key topics were called out to shape the quantitative topic guide and interviews

1. The potential and reality of Return on Investment for an e-Assessment system.
2. The increasing awareness and need for defensible exams utilising psychometrics.
3. The suitability of e-Assessment for the candidate population and to enhance the qualifications portfolio.
4. The use of multiple choice questions for higher levels of learning and high stakes exams.
5. Preferences on e-Assessment delivery test sites.
6. Technology Providers and the use of single technology platforms or multiple versions conforming to agreed standards.
7. The role of regulation.

### Number of Quantitative Respondents

93 respondents from 115 Awarding Bodies (currently accredited by QCA) responded to the quantitative phase of the research. This covers 81% of the research universe. This exceeds the 87 respondents from 116 Awarding Bodies from the previous study.



*38% of respondents are currently delivering an e-Assessment programme*

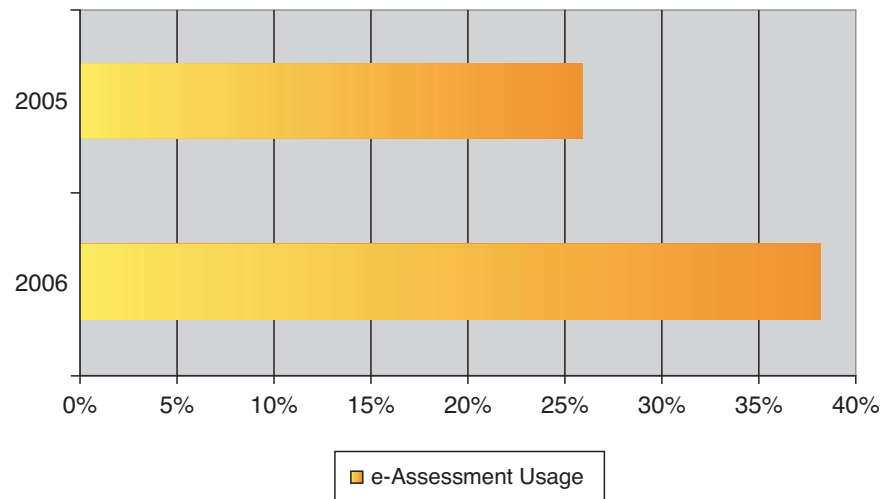
## Quantitative Results

### e-Assessment Delivery and Usage

**Q: Does your organisation currently deliver some form of e-Assessment? (Prompted)**

38% of respondents are delivering e-Assessment. This represents an increase from the 2005 percentage of 26% (46% year-on-year increase). Of the 62% who are not delivering e-Assessment, one-third are in the process of piloting or developing a system. Another third of this 62% believe that their qualifications are not transferable to an 'e' testing method.

Other reasons that were called out for non-adoption included the cost of e-Assessment, continuing to monitor developments and some Awarding Bodies had just relaunched their qualifications.



Base: 93 Respondents (35 running e-Assessment, 58 not running e-Assessment)



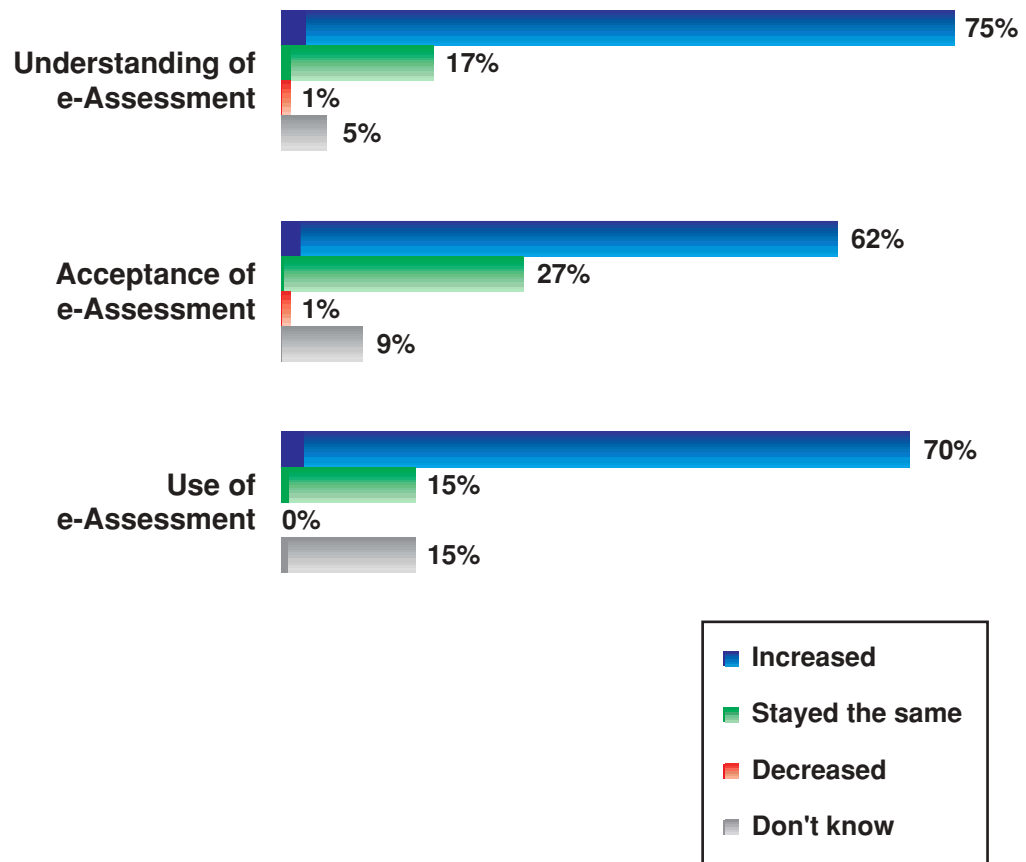
*The understanding, acceptance and usage of e-Assessment has increased in the last twelve months*

**Pace of change**

**Q. Over the last twelve months, would you say that the following have increased, decreased or stayed the same across the industry? (Prompted)**

1. Understanding of e-Assessment
2. Acceptance of e-Assessment as a valid method of testing
3. Use of e-Assessment

A strong majority of respondents believe that the understanding, acceptance and usage of e-Assessment has increased in the last twelve months. With more examples of e-Assessment being launched in the UK, the visibility of the e-Assessment market being called out by QCA and more mainstream media reports, this strong result indicates a quickening pace of change.



Base: 93 Respondents (35 running e-Assessment, 58 not running e-Assessment)



## Advantages

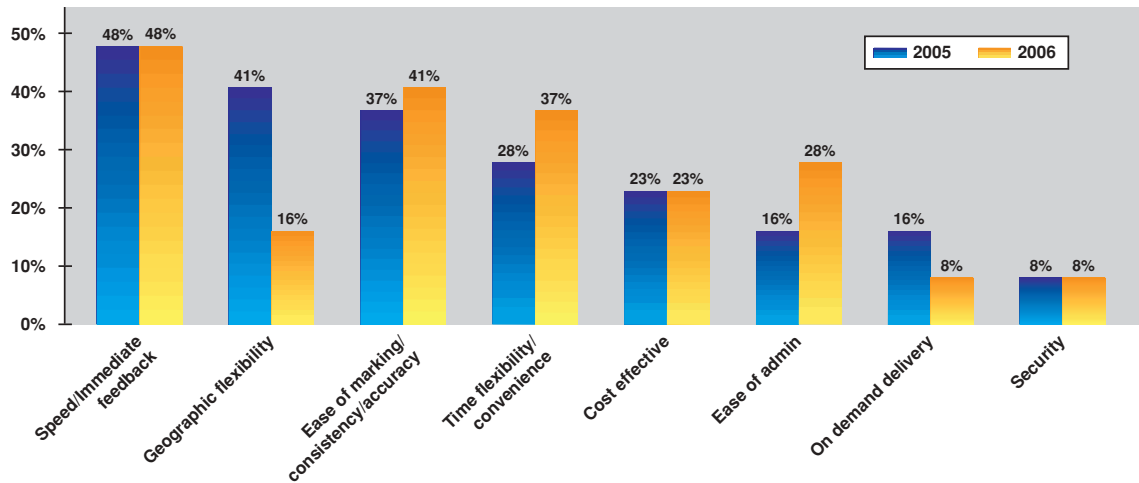
### Q. What are the three main advantages or benefits of using e-Assessment, in your opinion, overall? (Spontaneous)

*Immediate feedback, time flexibility and ease of marking are the current three leading advantages of e-Assessment*

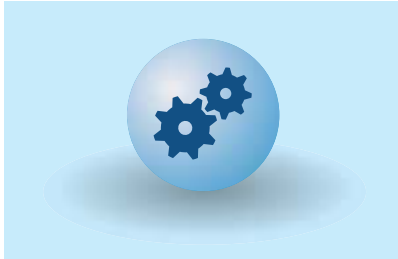
As Awarding Bodies start to use e-Assessment systems, it is interesting to note which advantages have emerged since adoption and which ones have not taken precedence. As uptake increases, the Awarding Bodies are now beginning to understand key benefits such as the ease of administration (increase from 16% in 2005 to 28%) and time flexibility/convenience (from 28% in 2005 to 37%).

It is clear that advantages such as geographical flexibility and on demand delivery are not 'front of mind' for Awarding Bodies as they were in the previous study. Business issues such as e-Assessment's ability to allow continuous testing from window/'burst' testing may be one factor in these changes.

### 2005 versus 2006



Base: 93 Respondents (35 running e-Assessment, 58 not running e-Assessment)



## Disadvantages

Q. What are the three main disadvantages of using e-Assessment, in your opinion, overall? (Spontaneous)

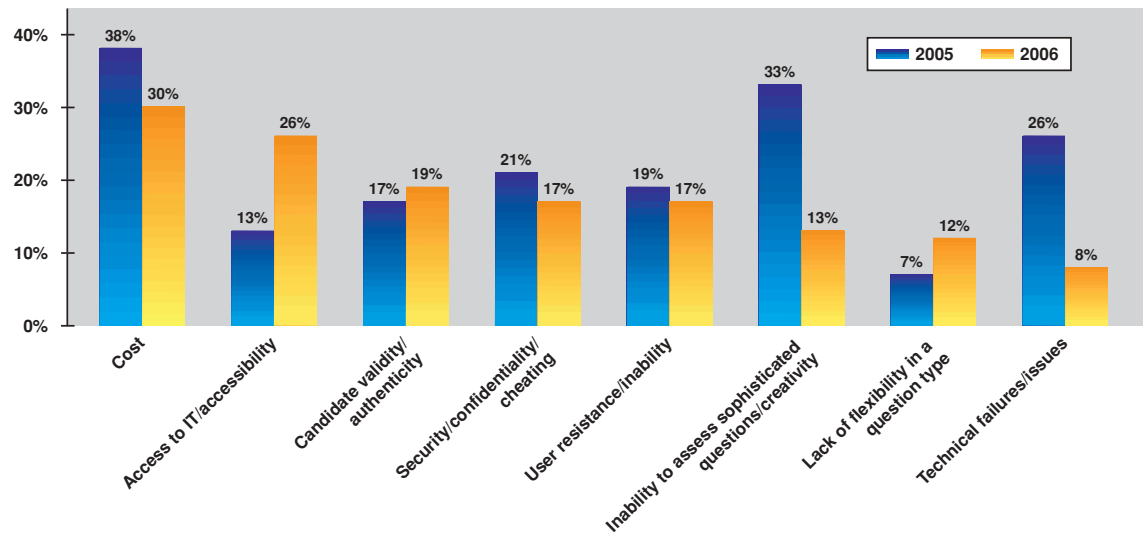
In terms of disadvantages, there are significant changes to the findings from the previous study.

For example, cost (down from 38% to 30%), the inability to assess sophisticated questions (33% to 13%) and technical issues (26% to 8%) have all decreased in importance as disadvantages.

This suggests that as the Awarding Bodies gain more familiarity with service providers and systems, there is a greater understanding and confidence in what e-Assessment can do.

*A maturing of the market and understanding of technology can explain the significant changes in these findings*

### 2005 versus 2006



Base: 93 Respondents (35 running e-Assessment, 58 not running e-Assessment)

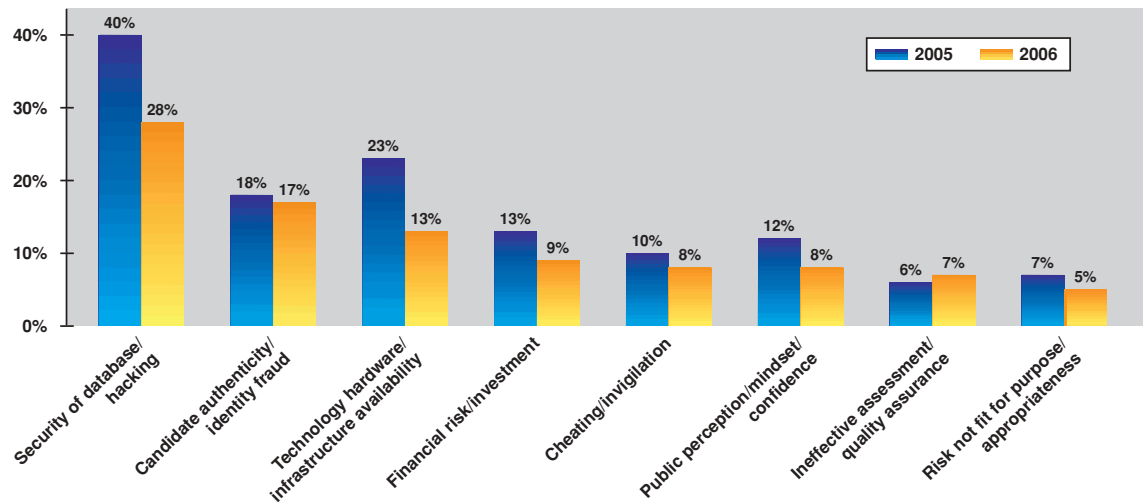


## e-Assessment Risks

**Q. What risks do you associate with the introduction of e-Assessment, if any? (Spontaneous)**

There are two key changes to the previous research findings. Both data security (fall from 40% in 2005 to 28% in 2006) and technology infrastructure (down from 23% to 13%) have fallen sharply. It is possible to suggest that as uptake of e-Assessment has increased, and maturity of common items such as IT hardware and data transmission/security systems have bedded down, more confidence in e-Assessment systems is being shown by the Awarding Bodies. Interestingly, the key risk area of candidate authenticity/identity has remained virtually unchanged, suggesting more effort and innovation is needed in this area from e-Assessment suppliers and Awarding Bodies.

*Fewer Awarding Bodies consider data security and technology to be risks of e-Assessment introduction*



Base: 93 Respondents (35 running e-Assessment, 58 not running e-Assessment)



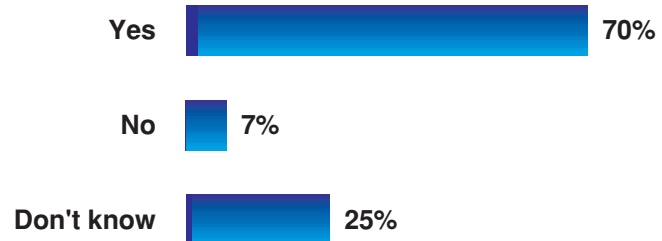
## Return on Investment

### Q. Do you believe that e-Assessment will deliver a Return on Investment?

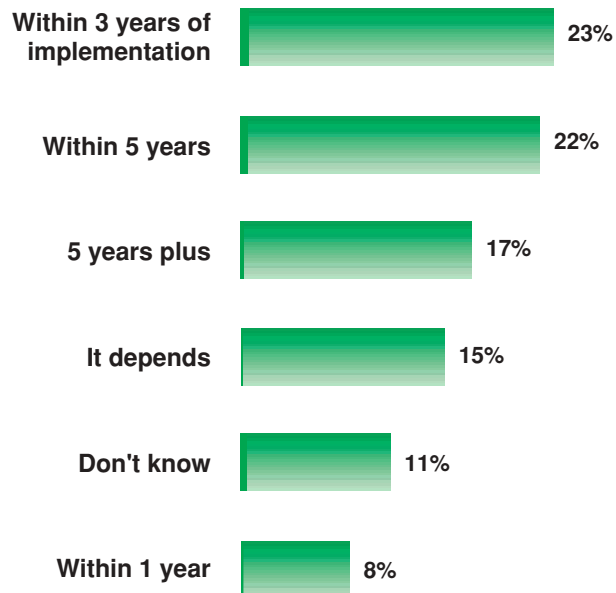
Seven out of ten Awarding Bodies believe that e-Assessment will deliver a return on their investment. 53% of Awarding Bodies believe that e-Assessment will show ROI within 5 years. Almost one-third believe it will deliver a return within 3 years.

Interestingly, only a small percentage (7%) said that it would not, with a quarter not able to give an opinion. This result strongly suggests that e-Assessment delivers a return on investment for UK Awarding Bodies.

*The results strongly suggest that Awarding Bodies believe e-Assessment delivers a return on their investment*



### Q. When do you believe you will, or did you, see ROI?



Base: 65 respondents who believe e-Assessment will deliver ROI



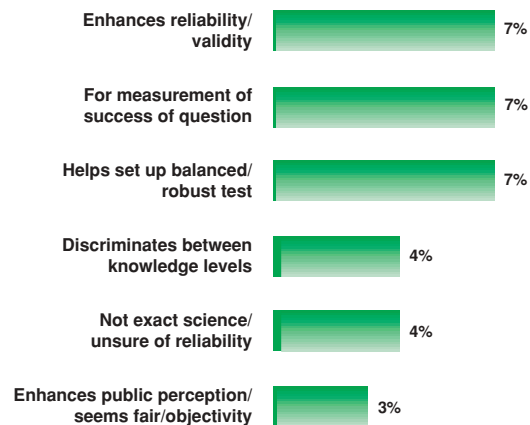
*Two-thirds of Awarding Bodies (68%) were unable to identify the implications or advantages of using psychometrics and over three-quarters of Awarding Bodies (77%) were unable to identify implications or advantages*

## Use of Psychometrics

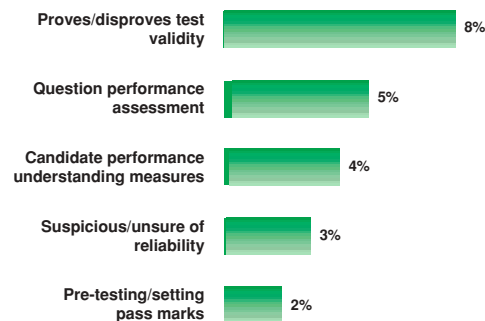
### Q. What do you understand to be the implications or advantages of using psychometrics to develop exam questions? (Spontaneous)

Whilst the issues regarding the use of psychometrics were flagged as a new emergent issue in the qualitative phase of the study, there was embryonic awareness within the quantitative phase.

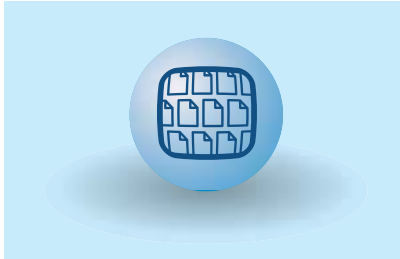
It is possible to argue that some of the larger Awarding Bodies who have statistician (or psychometric) staff members would be the respondents who would recognise the terminology and uses regarding psychometrics. The rate of growth in other areas of assessment knowledge since the last research study suggests that this issue will become more prevalent in a short space of time.



### Q: What do you understand to be the implications or advantages of using psychometrics to gain insight into candidate and exam performance? (Spontaneous)



Base: 93 Respondents (35 running e-Assessment, 58 not running e-Assessment)



## Driving Uptake

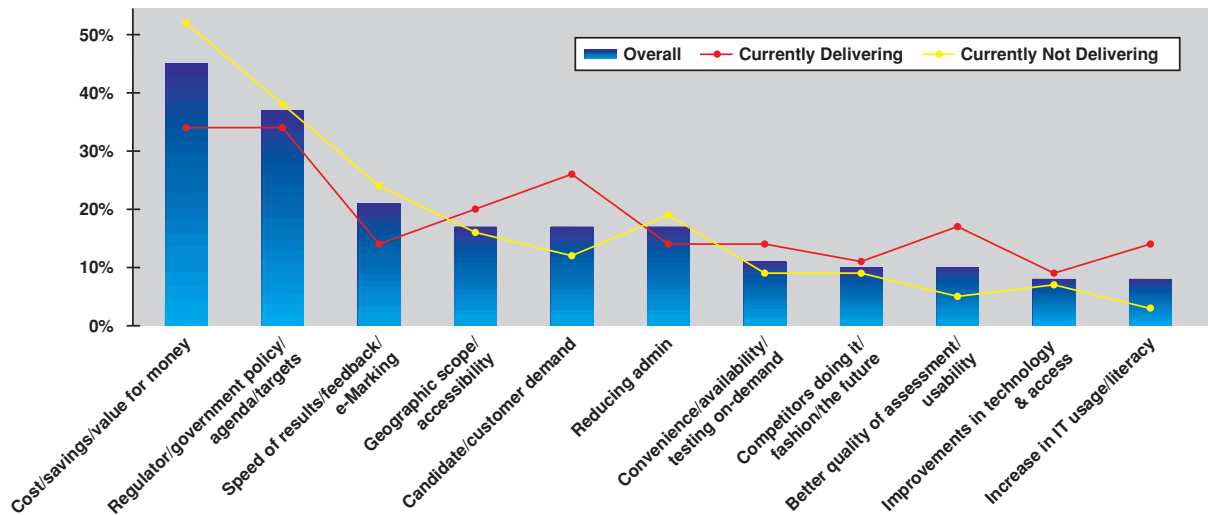
Q. Across the assessment industry, what do you believe to be the top three factors currently driving the uptake of e-Assessment? (Spontaneous)

In terms of first-recall, primary front-of-mind mention was made of the push from the regulator/government towards e-Assessment.

*Cost savings, regulator targets and efficiency are seen as the top three factors driving uptake.*

The overall current driver is the desire for cost savings, as it was in the previous study. Improving service standards for candidates, such as accessibility, convenience and speed of results are also key drivers.

Awarding Bodies currently delivering e-Assessment see factors such as candidate demand and a better quality of assessment driving demand more than those Awarding Bodies who do not deliver e-Assessment.



Base: 93 Respondents (35 running e-Assessment, 58 not running e-Assessment)

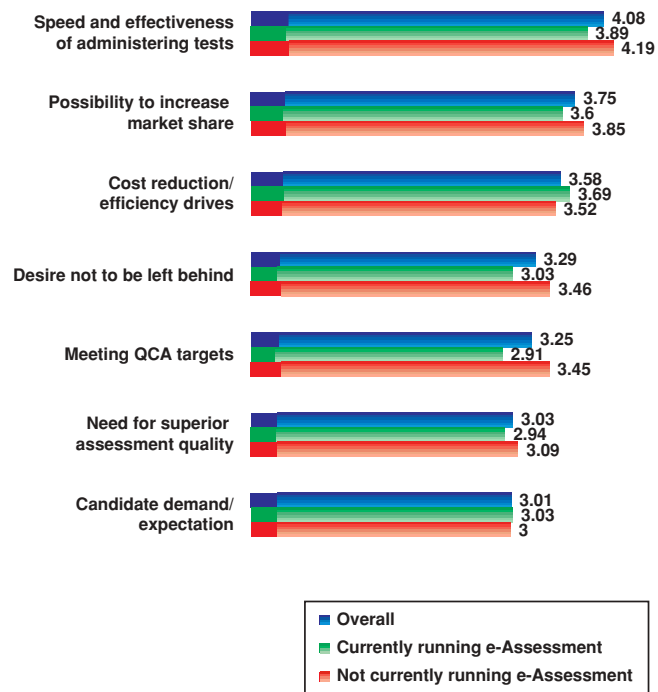


*Improving test administration is a key future driver for e-Assessment adoption*

**Q. How strongly, on a scale of 1 - 5 (where 1 is not at all and 5 is very strongly), do you agree that each of the following will further drive e-Assessment take-up rates over the next year? (Prompted)**

On a mean average basis, the desire to improve the speed and effectiveness of administering tests are considered to be the future drivers for e-Assessment take up.

Cost reduction is again considered to be a key future factor for those Awarding Bodies already running e-Assessment. Interestingly, those Awarding Bodies not delivering e-Assessment believe more in the possibility of increased market share that will drive up uptake. It is possible to argue that some of these organisations may therefore see e-Assessment as a differentiator in candidate choice.



Base: 93 Respondents (35 running e-Assessment, 58 not running e-Assessment)



### e-Assessment Importance Factors

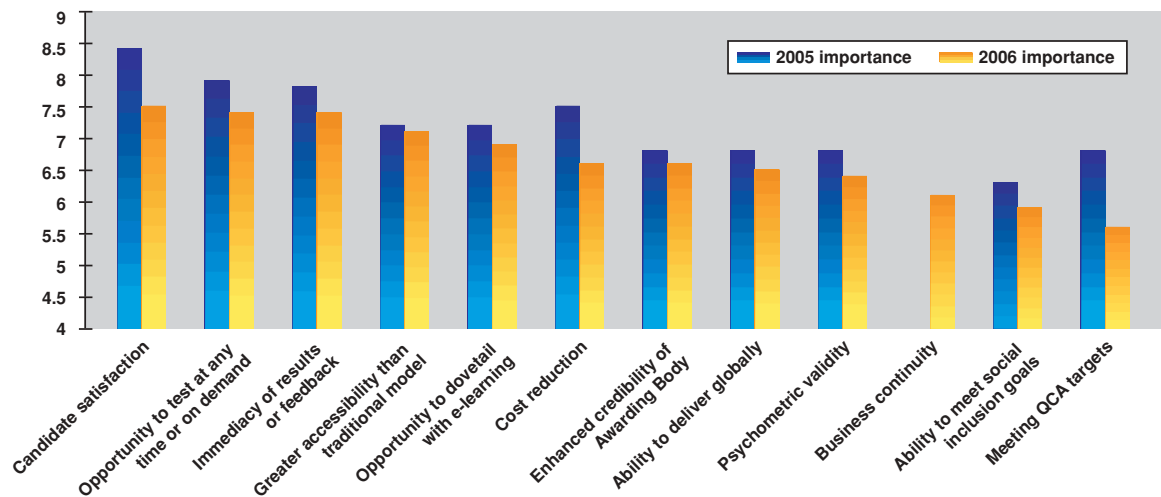
Q. On a scale of 1-10 where 10 is extremely important and 1 is not important at all, how important are each of the following factors to you in relation to e-Assessment? (Prompted)

Candidate satisfaction and 'on-demand' testing, with an average importance score of 8.4 and 7.9 out of 10 respectively are the two most important factors relating to e-Assessment for Awarding Bodies. Indeed, while they score highly in the previous study, these average scores are perhaps a reflection of e-Assessment adoption.

*Candidate satisfaction and the ability to 'test on demand' are the most important factors, and have grown in importance since the previous study*

The issue of business continuity, called out from the qualitative phase, is given similar importance by Awarding Bodies to meeting social inclusion goals and meeting regulator targets.

### 2005 versus 2006



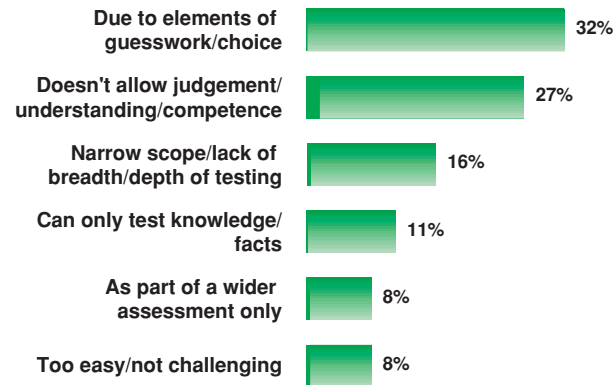
Base: 93 Respondents (35 running e-Assessment, 58 not running e-Assessment)



*The majority of Awarding Bodies (54%) believe that MCQs are suitable for high stakes exams and for higher level learning (52%)*

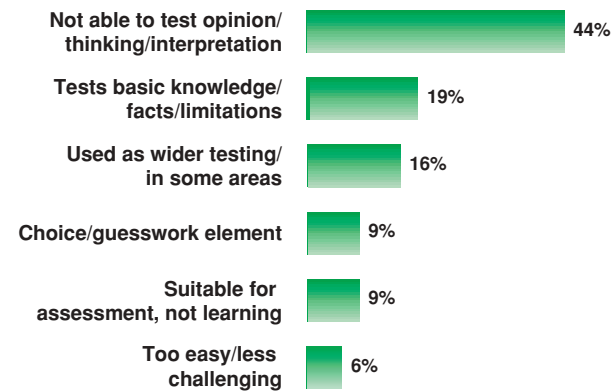
### Suitability of Multiple Choice Questions

Q. Do you believe that Multiple Choice Questions (MCQs) are suitable for high stakes exams? (Spontaneous)



Q. Do you believe that Multiple Choice Questions (MCQs) are suitable for higher level learning? (Spontaneous)

It is of great note that respondents answering yes for both questions were in the majority of respondents. Nevertheless, there is still a very significant minority that hold views that MCQs have a variety of key issues which prevent them being used for higher levels of learning and high stakes exams. An underlying notion that elements of guesswork play a part in diluting the effectiveness of MCQs is also a recurrent factor. It is possible to argue that the emergent visibility of psychometrics in the creation of exam questions and items will impact on these figures in the future.



Base: 93 Respondents (35 running e-Assessment, 58 not running e-Assessment)



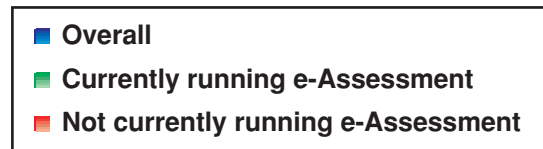
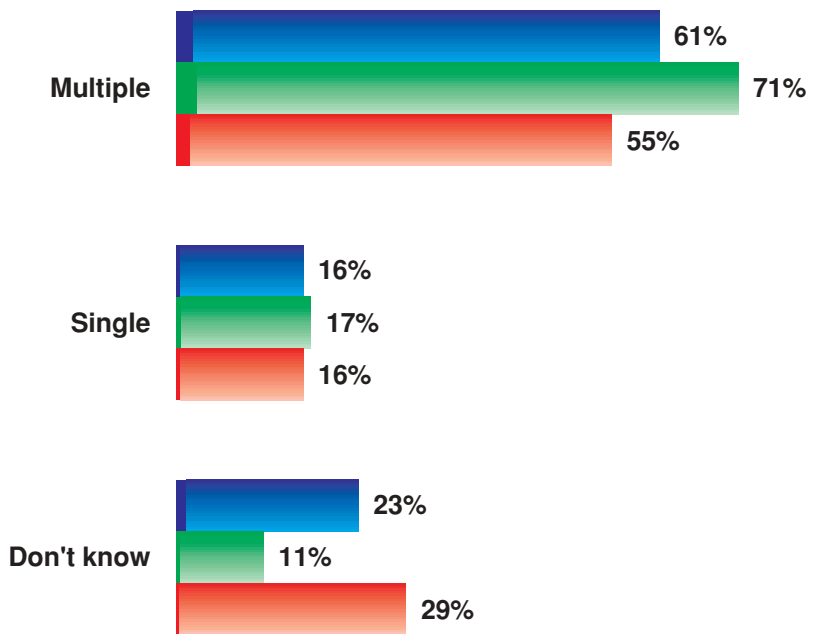
*Awarding Bodies prefer multiple e-Assessment technologies conforming to agreed standards*

### Technologies

Q. Is a single assessment technology preferable in the vocational/professional market, or multiple versions conforming to agreed standards? (Prompted)

A substantial majority of all respondents prefer multiple e-Assessment technologies conforming to agreed standards to a single assessment technology.

This is a clear call to organisations that want to adopt recognised e-Assessment standards such as QTI and BS7988. Awarding Bodies prefer these visible standards, rather than having to use a single assessment technology selected for them. This could be for reasons of innovation, business continuity, competition and continuous quality improvement.



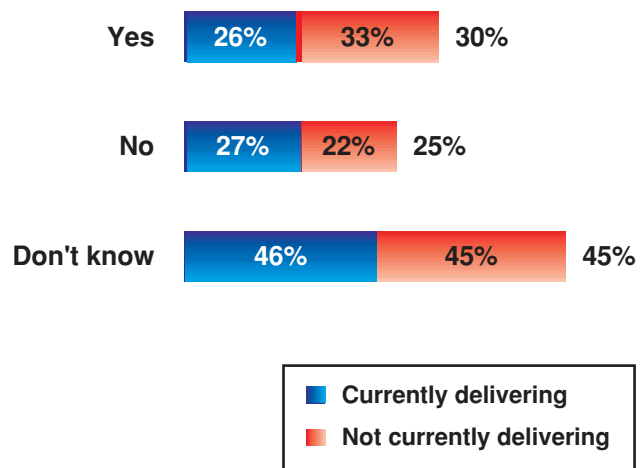
Base: 93 Respondents (35 running e-Assessment, 58 not running e-Assessment)



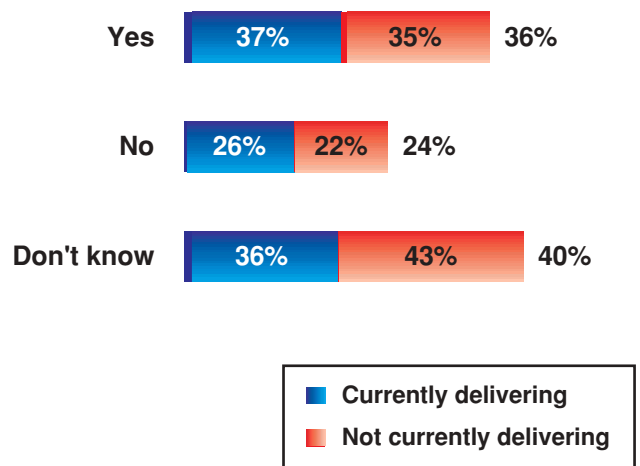
*The regulator's technical ability to assess e-Assessment capabilities/secureities and its understanding of transferring qualifications to e-Assessment is recognised by a minority of Awarding Bodies*

**Regulation**

**Q. Do you perceive QCA to have the competencies to assess the technical capabilities and securities of e-Assessment used by Awarding Bodies? (Prompted)**



**Q. Do you feel the QCA fully understands and supports the process for transferring qualifications to an e-Assessment methodology? (Prompted)**



Base: 93 Respondents (35 running e-Assessment, 58 not running e-Assessment)

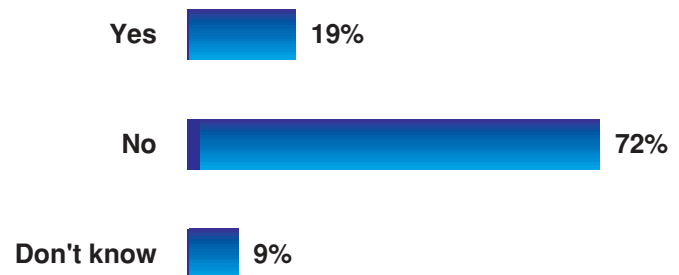


*Full adoption of e-Assessment by 2009 is not seen as achievable across the assessment market*

## Regulation

**Q. Do you believe that the QCA's vision of full e-Assessment adoption by 2009 is achievable across the assessment market? (Prompted)**

Agreement is down from 22% in 2005 research. 31% of those already running e-Assessment consider the Blueprint to be achievable, whereas 81% of those not yet delivering e-Assessment do not.



**Q: How satisfied are you with the FAB's (Federation of Awarding Bodies) activity to date in driving e-Assessment knowledge? (Prompted)**



It is important to note that just over 60% of respondents were not aware of FAB's involvement in driving e-Assessment knowledge and therefore were unable to answer the question.

Base: 93 Respondents (35 running e-Assessment, 58 not running e-Assessment)



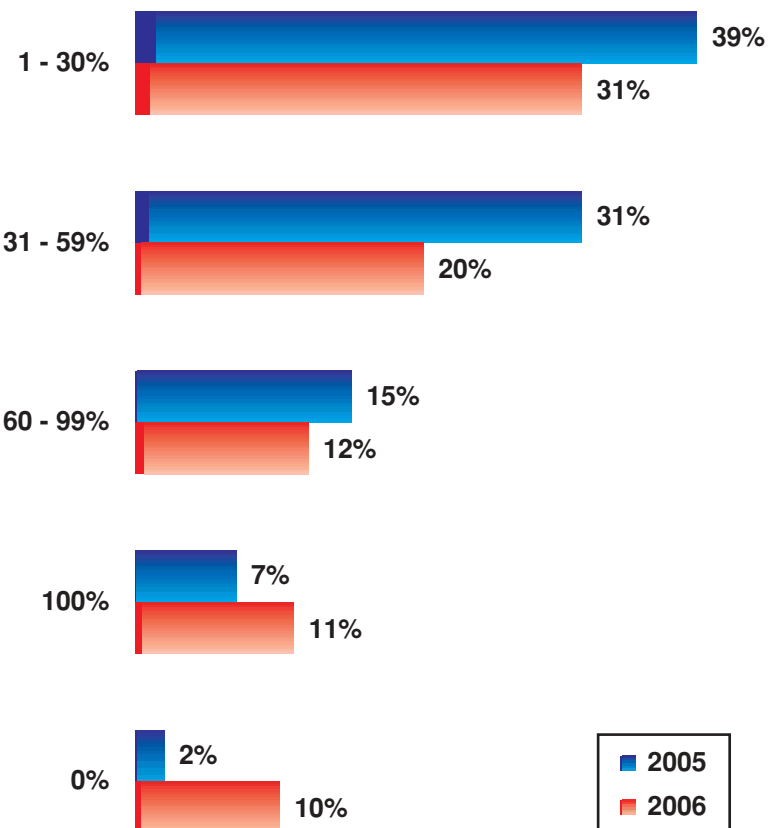
*Awarding Bodies are understanding the possibilities of what e-Assessment can do within their qualifications portfolio and for their candidates*

### Future of e-Assessment

**Q. Approximately what percentage of your assessment programme do you realistically expect to be e-Assessment by 2009? (Spontaneous)**

The maturing of the marketplace and bedding down of the technology may account for some of the changes since the previous study. A significant rise in those who completely dismiss the 2009 e-Assessment date (from 2% to 10%) is notable, although it is possible to suggest that this is down to lack of appropriateness, rather than undue pessimism.

Interestingly, more than one in ten respondents believe that 2009 is achievable for all of their assessment programme. This signifies clear confidence in their own investments and plans over the medium term.



Base: 93 Respondents (35 running e-Assessment, 58 not running e-Assessment)



*Issues such as Return on Investment are fully understood*

## Overview

The headline finding from this study is that 38% of Awarding Bodies surveyed currently use e-Assessment to deliver up to 60% of their assessment programme.

If the rate of change remains the same (as noted earlier), e-Assessment will soon be adopted by over 50% of Awarding Bodies: a clear majority. e-Assessment has now achieved 'acceptability' within the marketplace with strong majority verdicts on understanding, acceptance and usage.

The key benefits of current e-Assessment are now being understood as more organisations implement the changes. Market movement and increasing recognition of factors such as ease of administration and time flexibility are hallmarks of systems that have successfully been bedded into organisations and accepted by stakeholders. This is also relevant in relation to perceived e-Assessment risks, where issues that scored highly in the previous study (such as data security and technology availability) are now scoring lower.

An area of business concern that was raised in the qualitative phase was the notion of e-Assessment's return on investment (ROI). Clearly when significant resources are staked in e-Assessment, stakeholder interest in delivering organisational benefits are paramount. Seven out of ten respondents believe that e-Assessment will deliver ROI – clearly a sign of confidence in how it can improve not just the candidate experience, but also deliver efficiency savings and/or stakeholder value.

The subject of psychometrics was also flagged in the qualitative phase as being one of emergent, but increasing importance. Whilst there is limited understanding of the subject at large, this can be compared to the weak knowledge regarding item types that was highlighted in the previous study. If the positive results regarding multiple choice items is indicative of how the market can quickly assimilate e-Assessment knowledge, it would be reasonable to suggest that the market knowledge of the benefits of psychometrics in e-Assessment will rise quickly.

The focus on candidate needs is called out by a number of the findings. It is pleasing that whilst there is an acknowledgment of commercial factors, candidate needs such as accessibility and time flexibility remain at the forefront. The importance of candidate satisfaction remains a key importance factor for Awarding Bodies.

The need to make the most appropriate and best e-Assessment choice is a suggestion arising from the strong call-out for multiple technologies conforming to agreed standards. As wider issues such as the Unique Learner Number and ID cards impact on facets such as registration, exam booking and candidate verification, the need for differing e-Assessment systems to have a mutually compatible interface point, recognised and mature e-Assessment standards will become more important. Additionally, this raises a flag to e-Assessment providers to ensure that their systems are capable of adhering to the demands of these standards.

*Candidate demand and needs arising from e-Assessment are a key area of focus*

There is a minimal belief that the (often-quoted) date of 2009 as a target date for full e-Assessment is achievable. Given the rapid changes and uptake of e-Assessment systems since that date was suggested in the QCA's 'Blueprint for e-Assessment', this may influence contemporary thinking on an e-Agenda, covering e-Assessment and other linked areas that QCA has responsibility for.

## Conclusions

The acceptance and usage of e-Assessment has clearly grown at a substantial rate compared to the previous research study.

The strong confidence shown in e-Assessment's ability to deliver return on investment is a major finding of the research.

The use of psychometrics is emerging with some usage reported by Awarding Bodies. Multiple choice question usage for higher levels of learning and high stakes exams is more widely understood and acknowledged.

Disadvantages traditionally associated with e-Assessment such as cost and technical issues have decreased in importance as uptake has increased and technology has matured. In parallel, areas of risk previously thought to be inherent in e-Assessment (data security and technology in general) are not as prominent as areas which perhaps are not exclusive to e-Assessment. Candidate authenticity is a key issue called out in the findings.

The needs and desires of the learner/candidate continue to be at the forefront for organisations wishing to adopt e-Assessment or already using an incumbent system.

## Appendix

### Victorian Curriculum and Assessment Authority of Australia Roundtable - October 24 2005

#### Strategy, Technology and Assessment Ken Boston - CEO QCA

In the world of education, technology for assessment and reporting is the third of three potentially transformative but still incomplete major reforms in the procurement of goods and services by large education systems in the last twenty years.

One is the private funding of public infrastructure, or PFI. Education systems have variously turned to the private financial sector to build, service and even operate schools and colleges, with the return on investment being generated by multiple use of the facilities on a commercial basis.

Outsourcing of the provision of goods and services is the second: maintenance, security, courier services and so on. In these remarks, I want to focus on maximising the opportunities technology offers, ameliorating the risks and on the strategies that governments and organisations might find most effective in achieving the benefits of reform.

I begin with the market-place of system-wide assessment and reports.

#### 1. On-screen testing

- *The replication of a paper test on a computer screen. This is a 'quick win' with some benefits. It requires no change in test design or the setting of questions, yet it eliminates the printing and physical distribution of test papers, and it allows location flexibility and time of testing.*
- *Or you can buy tests which, like paper tests, use closed tasks or questions for which there is a known and pre-specified right answer, but which also use media clips, drag and drop actions, oral language, or prompts and clues-modifications which are not replicable on paper.*
- *Transformational on-screen tests exploit interactivity capabilities between IT, data and media, using scenario-based environments and a 'virtual world' data set, and include open tasks that have no predetermined answer or solution. Candidates demonstrate their ability through process, as well as producing a range of outcomes.*

## 2. Managing test questions or test items (three different solutions)

- *For tests to be available beyond a one-off bounded time slot, you need an item bank from which tests and tasks can be selected, to avoid the potential for duplication.*
- *You need randomised selection or version control, so that organisations and candidates cannot predict the content of the test - something new and different always comes out of the item bank.*
- *If you want to move away from simple paper translation, you need to develop new types of question items and tasks. New skills are required for item development with modifications not replicable on paper, and especially for the development of items for transformational on-screen testing. And that is because the question-setting role of the examiner, and the software development role of the supplier of technical solutions, have merged to become one activity.*

## 3. Coursework processing/management as part of an overall assessment in a subject

- *This consists of the submission and marking of coursework submitted in computer or multi-media format, and the issuing of results. It can include digital artwork pictures, sound, video presentations etc.*
- *You can also buy ongoing e-portfolio management, enabling access to the e-portfolio by assessors and verifiers, and maintenance of a personal learning record.*

## 4. On-screen marking

The particular skill is to scan written answers electronically, present the answers on computer to external markers; and to gather and process the results. The product is strong and convincing:

- *better quality marking, through early detection and remediation of aberrant marking;*
- *random distribution of scripts and items to markers;*
- *specialisation of markers in a limited number of items;*
- *reduction of clerical errors, because the computer sums the marks;*
- *elimination of paper distribution; and*
- *greater security.*

## 5. Automated marking (three versions)

- *The simplest and widely used is automated, multiple choice marking done by machine.*
- *The second is automated short-answer marking, using character recognition software.*
- *The third is process-based marking. It requires increasing intelligence in the marking engine and redesign of the tests to make the tasks and responses more amenable to automated marking electronic capture of candidates' responses.*

## 6. Data capture, processing and reporting

- *You can buy systems which hold data on all units, subjects and qualifications in your qualifications framework, on prerequisites, syllabuses, assessment methods etc.*
- *There are systems which allow learners to assemble e-transcripts of their results and qualifications, across the various agencies responsible for the assessment, awarding or accreditation of its components.*
- *And there are systems of data capture and reporting which generate data in forms useable by governments and Awarding Bodies to monitor and report performance, to allocate resources, to plan interventions, and to set targets for the future.*

### **Regulation and the demand side**

The purpose of the regulation of assessment is essentially to protect the interests of the learner - the interests of the demand side - in matters of the maintenance of standards, the fairness of the examinations, the quality of the marking, the quality of the grading, and the national and international standing of the credential which is the outcome.

In England, the separation of the demand side from the supply side is clear. The syllabuses, assessment methods, marking and grading processes and issuing of results for the general qualifications are the responsibility of the 'big five' major UK Awarding Bodies, monitored by the QCA. The UK Awarding Bodies have the interests of learners at heart, but they are squarely part of the supply side; the demand side champions are the QCA and the other regulatory authorities.

### **Security of the supply chain**

The effect of technology is not to make the supply chain run more smoothly, but to change it. Much effort in the current system is focused on the post-test stages such as marking, moderation, re-marking, grading and the processing of results; there is less upfront in test development. While not downplaying the critical importance of test development and pre-testing, most of the effort in terms of hours and labour is squeezed into the last two months.

Technology changes this. Both automated marking and on-screen marking reduce the major 'back-end' logistical effort in managing answer sheets and scripts; automated marking also removes the dependence on high volumes of markers, and removes the processes of marker recruitment, standardisation and grading.

At the same time, there is much more 'front-end' work: pre-testing, analysis of pre-test responses, production of the test items, and establishing the quality, standards and comparability of individual items. It is necessary to know how each item will perform before the test goes 'live'; the items are grade related, and there will be no opportunity for their performance to be adjusted later through an awarding or grading process.

Further, there will be increased and alternative routes through the system, depending on which e-assessment components a qualification uses. For example, one assessment may comprise four different elements: multiple choice, short answer, voice capture and essay-type responses, each of which may follow different automated and non-automated marking routes. Although each individual route might be simpler than the paper-based system, there is more complexity overall because of the combination of e-assessment components and the increased process options.

The present simplicity of a contracted supplier providing a specific good or service to an Awarding Body or assessment authority is thus much harder to maintain in the new world of technology. The e-assessment system partners will include e-test design suppliers, e-test service providers, e-portfolio service providers, technical support providers, scanning bureaus and testing centres - all jointly needing to be involved in many aspects of test design, usability, marking design and issues of comparability. Actions of all these players have the potential to impact on the final assessment for candidates. As the number of organisations and the complexity of their actions and responsibilities increases, how is government to ensure that it all works in the interests of the demand side - that is, that it ensures a fair deal for the learner? Clearly, there needs to be an e-Code of Practice.

This is widely recognised in the UK. There is work being done on an emerging, voluntary Question and Test Interoperability Standard (QTI); there is BS7988, which sets out some minimum requirements for organisations that use computers to make assessments; and there are standards regarding data security.

#### **Quality assuring the ether**

The notion of a national delivery system raises the second issue relating to regulation and quality assurance that needs to be considered in the procurement exercise. Can there be such a thing as a national delivery system in an environment in which e-learning and e-assessments are available electronically anywhere on the planet? How can one quality assure the ether?

Obviously, there cannot be national delivery systems, in the sense of monopolistic provision. The internet makes that impossible. But there can and must be regulation in the interests of quality assurance. There is potential in examining whether the appropriate strategy for that might be through cooperation by groups of nations such as OECD countries, to use their own national qualifications frameworks to that end.

An appropriate strategy for e-regulation might be for each country to include in its qualifications framework those qualifications which have content and assessments originating from outside its jurisdiction, and which have been quality assured by the country itself, or by a partner country which is recognised as having similarly robust quality assurance procedures. Thus, qualifications which are currently distinctively English or Australian or German might be taught electronically and assessed electronically in a range of countries, along with a vastly increased range of vendor qualifications.

#### **Transformational on-screen testing**

It is difficult to compare the results of transformational tests with existing tests. They measure different things in different ways. Their development therefore will be directed to the assessment of new qualifications, or to assessing a current syllabus in a new and different way. Time-series comparisons of performance cannot continue once a test of this type has been introduced.

The new technology makes it possible to evaluate skills and knowledge not possible with existing paper and pencil tests.

### **On-screen marking**

On-screen marking is in one sense the simplest and most benign of technological interventions. So, during the marking season, markers only need a personal computer. No longer bundles of scripts be delivered by post; the electronic distribution of written responses to questions is random; and markers deal with responses to particular questions, rather than mark each script as a whole. The latter means that the mark scheme for each item is followed more closely, because there is no capacity for a marker to be influenced by the global 'feel' of the script as a whole.

Therein lies an issue. On-screen marking against specified mark schemes, with early detection and correction of aberrant marking, greatly increased second marking, and the automatic addition of marks assigned to individual components of assessment, is likely to produce a result which is more valid and reliable than manual marking, except in situations where second examining and close supervision by senior markers is just as intense.

### **Offshore marking**

One of the properties of the new technology is that it can use the 24 hours in the day. Up till now, marking has been after work hours, sometimes during work hours, at weekends and in holidays. When scripts and marks are transmitted electronically, they can be marked around the clock.

Are we ready for that? There are clear financial imperatives for providers of e-assessments to operate internationally. The investment required to support solutions is huge. Much of the innovation and assessment we are looking for will not be possible unless such investment is made. But it is also clear that no country can permit its marking to go offshore, unless it can account for the change in terms of it improving the quality of marking and thus being in the interests of the learner.

Our various jurisdictions therefore need strategies to explain to the public why it is of benefit to the learner that papers previously marked at home might now be partly marked overseas. In some countries a clear incentive is insufficient availability of local markers for the periods when they are required. It is patently in the interests of learners that marking is as accurate and reliable as possible; it is axiomatic that quality depends upon early detection and correction of poor marking; and it is clear that working electronically around the clock maximises the time which can be given to the marking process within the short national time frames that are generally the rule for completion of marking. The emphasis must be placed on the guarantee of quality rather than early return of results.

### **Change in the nature of the tests**

It is commonly said that the increased opportunities for on-demand or when-ready testing will create a tendency for the delivery of smaller and smaller units of learning and assessment, and that it will therefore be imperative to test synoptically across whole sets of units to evaluate overall knowledge, competence and understanding. Similarly, the point is frequently made that the technology must serve the purpose of the test, and that there is a danger that in striving to make the tests computer based, some important objectives will be lost.

### **Is it worth the candle?**

Do the risks significantly outweigh the opportunities? Is it all worth the effort?

All of these involve changes in existing processes rather than the embracing of a new assessment and reporting paradigm. Together, they raise risk questions about the security of the current assessment regime, once it becomes electronic; about the need for an electronic rather than paper-based Code of Practice; about how governments might explain the inevitable blip in the time series of national performance in a subject; about offshore marking; about the recruitment and training of markers; and about whether the tests are being developed to fit the technology rather than the technology to fit the tests. They also raise the question of whether there is any real long-term benefit in automating imperfect processes and imperfect assessments.

Most risk lies in the apparently easiest solutions.

There is no political downside in evaluating skills and knowledge not possible with existing pencil and paper tests. Nor, provided such innovations are successfully piloted over time periods determined by the technical requirements and the provision of the necessary resources, is there any fundamental additional requirement for implementation other than high level management skill. I conclude by saying that all of us have some immediate needs for the improvement of existing assessment and reporting, and managing the risk. But the real prize is in the other basket: the transformation of assessment and reporting, in the service of profoundly better education.

Is it worth the candle? Absolutely.

**The Thomson Corporation and Thomson Prometric**

The Thomson Corporation, with 2005 revenues of \$8.70 billion, is a global leader in providing integrated information solutions to business and professional customers. With operational headquarters in Stamford, USA, Thomson has approximately 40,000 employees and provides services in approximately 130 countries.

Thomson Prometric is the recognised global leader in technology-enabled assessment services, providing paper-and-pencil, Internet and computer-based testing solutions. It offers a fully integrated testing system that includes test development, test delivery and data management capabilities. On behalf of 600 current clients in the academic, professional, government, corporate and information technology markets, Thomson Prometric develops and/or delivers assessments through a global network of testing centres in 132 countries as well as direct to candidates via the Web.

## References and Further Reading

- Boston K, (2005) Strategy, Technology and Assessment
- Eccleston, K (2005) Understanding Assessment and Qualifications in Post-Compulsory Education and Training. NIACE 2005 (2nd Ed.)
- Haladyna (2004) Developing and Validating Multiple-Choice Test Items (3rd Ed.)
- Harden (1999) Constructing Multiple Choice Questions of the Multiple True/False Type
- Mills, Potenza, Framer, Ward (Ed.) (2002) Computer-Based Testing - Building the Foundation for Future Assessments
- Osterlind (1998) Constructing Test Items - Multiple-Choice, Constructed-Response, Performance and Other Formats (2nd Ed.)
- Parshall, Spray, Kalhon & Davey, (2002) Practical Consideration in Computer-Based Testing
- PriceWaterhouseCoopers (2005) The Market for Qualifications in the UK

For more information regarding this research, please contact [ukmarketresearch@thomson.com](mailto:ukmarketresearch@thomson.com)

Copyright 2006 Thomson Prometric. All rights reserved. No part of the material protected by this copyright may be reproduced or utilised in any form or by any means, electronic or mechanical, including photocopying, recording, and broadcasting, or by any information storage and retrieval system, without permission in writing from Thomson Prometric.

Thomson, Thomson Prometric, and the Thomson logo, are registered trademarks of The Thomson Corporation and its affiliated businesses. All other trademarks, registered trademarks and service marks are the property of their respective owners.