The future of assessment

Tim Oates CBE| *Group Director of ARD* **| 03 June 17**

It's all about technology…

…technological development is only a small part of well-evidenced, coherent development of assessment…

It's all about technology…

On-screen marking – remote standardisation and moderation

Marker monitoring

On-screen assessment

Automated marking

E-portfolios

Item banking

Curation of all our questions

On-demand testing – versioning, time zones

Adaptive testing and automated test construction

Rapid feedback

Results analysis – us, schools, candidates

Test evaluation and development

Experimental work – rank ordering, on-screen marking environment

Formative assessment environments and assessment embedded in learning

Cambridge Assessment

It's all about technology…

Learners

Users

The content of learning

The accuracy and fairness of measurement

Access to assessment

Reporting

Impact

Cambridge Assessment

Validity

Using an elaborated 'consequentialist' model

Uses - impact

Working upwards from the construct

Assessment in service of curriculum aims

Evidence on key constructs

Criteria relating to assessment – 'Cambridge Approach'

Attending to the **PURPOSES OF ASSESSMENT**

Reliable
consistent measurement

Valid
measures precisely what it claims to measure

Sound construct base
measures something consistent with curriculum aims

Consequential validity
the uses to which the assessment is put are technically and ethically sound

Beneficial impact
the full range of effects are beneficial

Utility
cost, resource

Cambridge Assessment

In assessment, the concept of *'construct'* is vital

Can multiply two three digit numbers

Understands and is inventive with metaphor

Reads a wide range of books for pleasure

Understands diffusion across a membrane

Can understand and use familiar everyday expressions and very basic phrases aimed at the satisfaction of needs of a concrete type

Understand the concept of percentage and calculate pc

Use the concept of inequality to analyse social relations

Understands conservation of mass

Measures accurately to quantify oxidation

Has successfully converted to Alouette III

Is intellectually well prepared for pre-clinical medical education

Has a certain level of verbal reasoning

Manifests externalising behaviour

Cambridge Assessment

What about '21st Century Skills'?

21st Century skills: Ancient, ubiquitous, enigmatic?

Irenka Suto Research Division Cambridge Assessment

Paper published in January 2013 in

Research Matters: A Cambridge Assessment Publication
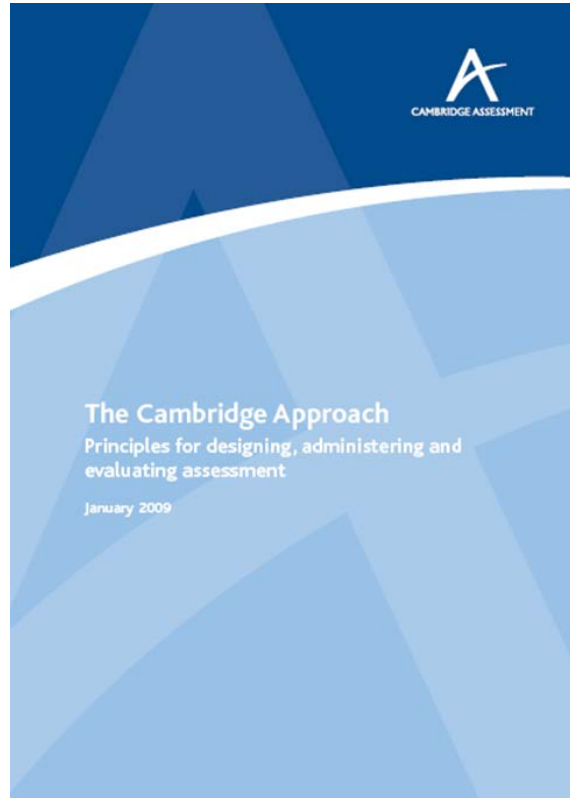
Contact details:

Dr Irenka Suto

Principal Research Officer

Research Division

Cambridge Assessment

1 Hills Rd, Cambridge, CB1 2EU

E-mail: suto.i@cambridgeassessment.org.uk

The Cambridge Approach
Principles for designing, administering and evaluating assessment

January 2009

http://www.cambridgeassessment.org.uk/Images/109848-cambridge-approach.pdf

Course categorisation

'Traditional' essay questions

Such as three-hour essay papers, intensive marking by specialists. There is increasing interest in the development of intelligent systems for automated marking

Short answer items

Typically groups of items focussed on dimensions of a cognate area; ARD is undertaking work on the comparative performance of specialist and non-specialist markers

Objective response items

Where the candidate has to provide a highly defined response, such as adding a line to a diagram, completing a number sequence, etc

Cambridge Assessment

Course categorisation

Multiple choice items

Typically a stimulus followed by a four-option response format; Increasingly criticised by the education establishment as 'reductionist', but sophisticated and well-trialled multiple choice items have a role to play in efficient, valid testing

Coursework and teacher assessment

Structured activities typically marked by teachers and tutors who are also responsible for managing the learning programmes in which the coursework is located

Performance assessment

Critical to vocational qualifications; outcomes- and competence-based assessment by qualified assessors. Guided by occupational standards – assessment based on direct observation of performance and collection of evidence into portfolios

Cambridge Assessment

Three key development and management models

- Banked items – 'data accumulation'
- Awarding-based assessments – 'review and reward'
- pPerformance-based assessment – 'evaluation'

Three key development and management models

1) Banked items – eg Cambridge English tests

design – validation – pre-testing – banking (with meta data) – administration – outcomes – [data accumulation]

Where items are subjected to pre-testing in order to yield item performance data, allowing inclusion in the bank when quality criteria have been met. The administration of the items does not include an awarding session, and yields more data which is accumulated in the record for the item in the bank. Secure administration of the tests allows re-use and data accumulation. The standard evaluation method for this model is integrated within the normal test administration process through the data-accumulation processes. Additional reviews of use and impact are commissioned as separate processes.

Cambridge Assessment

Three key development and management models

2) Awarding-based assessments – eg GCSE, PreU

design – review – administration – outcomes - awarding - [review]

Where items are not subjected to pre-testing, and the total score in an examination or examination component is scrutinised in awarding processes, designed to align the standard in any session with the standard applied in earlier sessions. Internal quality assurance processes operate on marking and awarding, using data derived from the awarding session, previous sessions and other bodies of data such as prior attainment of candidates. The main evaluation process is through separate review and special processes such as comparability exercises.

Three key development and management models

3) Performance-based assessment – eg OCR vocational awards

specification of standard – identification of opportunities to collect evidence – judgement – verification – [evaluation]

Where assessment is based on simulated or naturally-occurring activities which allow assessment against stated standards. Judgement of outcomes is made by competent assessors against those standards. Samples of evidence are subjected to external verification. Overall evaluation is discharged through separate review.

The Cambridge Approach

The integrated model

All assessments originated and operated by the Group, are underpinned by an integrated model of design, development, administration and evaluation. Fundamental to this is:

1) Clear statement of purpose
2) Identification of candidate population and their characteristics
3) Accurate cataloguing of the constructs which are the focus of the assessment
4) Clear statement of measurement principles, including standards-setting, standards monitoring and standards maintenance processes
5) Documenting of development and refinement processes
6) Documentation of all elements of administration including provision of clear service standards for each assessment
7) Documentation of all quality assurance processes
8) Analysis of the measurement characteristics of the assessment, including its reliability; analysis for bias; and analysis of other aspects of the technical performance of the assessment
9) Analysis of operation of the assessment, its impact and the use to which outcomes are put
10) Documentation of all key decisions which affect the characteristics, operation and use of the assessment and its outcomes, and archiving of relevant material

Cambridge Assessment

Observation and data

Learners – data (prior attainment, concurrent attainment; patterns; trends)

Users – feedback; predictive validity

The content of learning – construct specification; curriculum impact

The accuracy and fairness of measurement – item design; marking; marker monitoring awarding; standards-setting, standards-maintenance; standards-monitoring

Access to assessment – utility; special arrangements; mode

Reporting – feedback granuality; washback into learning; feedback into instructional design

Impact – wider evaluation of cost, quality, curriculum function etc

Cambridge
Assessment

Standard-setting – international benchmarking

Standards maintenance – comparability

Standards monitoring – reference tests

Stats

Versus

Judgement

[comparable outcomes; script comparison; rank ordering & structured judgement; benchmark centres; syllabus pairs; concurrent attainment]

Challenges and changes

Technology in assessment – delivery, transformation (eg adaptive testing)

Political aspirations and tendency to use quals for all changes

Social demands for transparency and accountability – trust

Increases in regulation – control

Developments in society and economy - aspirations

International development and economic changes

Cambridge Assessment

Provision of examination services

Provision of educational services

The seamless integration of assessment and learning

Establishing high quality formative assessment

Using formative assessment for summative purposes

Managing the functions of qualifications – ensuring beneficial impact of assessment

The growth of national regulation

The structure of international markets in assessment

Harnessing the benefits of competitive arrangements

Endless demand for certification?
Assessing everything which moves

Cambridge
Assessment

Influencing thinking

Supporting operations

Contributing to fundamental knowledge

Increasing professionalism in assessment

Improving public confidence in assessment

Q: *Stop the Train*

A steam train moving with a speed of $v_0 = 60 \, \text{km} \, \text{h}^{-1}$ is brought to rest by first shutting off the steam, when the train runs against a resistance equal to $1/100$ of its weight, and later by applying the brakes, at which point the train runs against a force equal to $1/8$ of the weight of the train.

| **Answer Now** | Hint 1 | Hint 2 | Hint 3 | Hint 4 | Hint 5 |

If the steam is shut off when the train is a distance $d = 0.33 \, \text{km}$ from a station, find the distance from the station that the brakes must be applied in order that the train may be brought to rest in the station.

Value

?

Please answer to an appropriate number of significant figures.

Units

Please choose an appropriate unit of measurement.

Don't forget to use the **hint** tabs above if you need help.

Check my answer

We can enhance assessment

We can improve equity, attainment, enjoyment, progression and labour market mobility

Benefits to individual, society and the economy

Cambridge
Assessment

Tim Oates CBE
Group Director of ARD