

# Assessment instruments over time

Gill Elliott, Milja Curcin, Nat Johnson, Tom Bramley, Jo Ireland, Tim Gill and Beth Black Research Division

First published in *Research Matters*, Issue 7, January 2009

## Introduction

As Cambridge Assessment celebrated its 150th anniversary in 2008 members of the Evaluation & Psychometrics Team looked back at question papers over the years. Details of the question papers and examples of questions were used to illustrate the development of seven different subjects. In each case the following research questions were addressed:

- Has the assessment structure altered over time?
- Have the emphases on different topic areas changed over the years?

The seven subjects studied were:

Mathematics	Physics	Geography	Art
French	Cookery	English Literature	

## Background

In the 150 years since Cambridge Assessment/University of Cambridge Local Examination Syndicate has been in existence, there have been a great many educational and social changes affecting students, teachers and assessments. This project sought to describe some of these changes and to illustrate them through changes in question papers. The project was a departure from the usual qualitative and quantitative methods used by the Evaluation Team, and instead took the form of a semi-structured investigation of the development of a number of subjects through the questions presented in the written examination papers.

These studies cannot be used to provide a commentary on *standards* over time, for several reasons:

- First, they do not contain sufficient salient information about the mark schemes, the curriculum and the exact nature of the work produced in response to the questions (scripts). Without *all* of these pieces of information, most of which no longer exist, comparisons about whether a particular era is 'better' simply cannot be made.
- Secondly, examination questions have changed over the years. For example, advances in technology have made it possible to routinely calculate statistics about questions (e.g. facility values) which can provide question writers with important feedback about the performance of that question. Additionally, much development has occurred around question writing and question writer training. Older questions which may seem difficult to 21st century readers may have been difficult for reasons which would nowadays be challenged on the grounds of fairness or validity. Finally, the regulation and oversight of all Awarding Bodies has changed beyond recognition in 150 years. Therefore, simplistically comparing questions from one era with another as evidence of changes in standards over time is flawed.

- Thirdly, the nature of the cohort has altered over the years and examination questions do not show this. So for example, the candidates sitting a School Certificate examination in 1907 might have been only a tiny proportion of the 16-year-old population, whereas the vast majority of 16-year-olds enter for GCSEs in the current context. As a consequence the level of accessibility of the questions differs – modern questions must be worded in such a way that all students being targeted can make some attempt at answering. The target candidature of past questions (particularly those from the earliest years sampled) was undoubtedly very different.

However, studies such as these can be used to illustrate the vast changes that have occurred, and the examples which follow show a small selection of the findings in each subject. These were presented as a poster at the 34th International Association for Educational Assessment (IAEA) Annual Conference which was hosted in Cambridge from 7–12 September 2008 by Cambridge Assessment, as part of the celebrations for its 150th anniversary.

The studies looked at the way in which papers were structured over the years, as shown in these examples from the **Physics** study (Table 1).

Table 1

Year	Paper	Time	Rubric	Example question
1927	Physics I	2hrs	Not more than <b>six</b> questions are to be attempted.	Explain the phenomenon of dew, and discuss the conditions which favour its formation. How is the dew point determined, and how can the relative humidity of the atmosphere be calculated when the dew point is known?
1957	Physics Ordinary Level Theoretical Paper	2½ hrs	Answer <b>all</b> the questions in Part I and <b>five</b> questions from Part II including at least <b>one</b> question from each of the Sections A, B, C.	{From Part I}: What is the freezing-point of water on the Fahrenheit scale? Express, in °C, a temperature which is 45 degrees below the freezing-point of water on the Fahrenheit scale.
2007	1982/4 Science: Physics extension option A Paper 4 Higher Tier	45mins	Wide range of mark totals per question	This question is about generating electricity. In 2005 the Prime Minister, Tony Blair, called for a 'National Debate' on nuclear power, climate change, and renewable energy sources. (a) Explain what is meant by a <b>renewable energy source</b> . [2] (b) More nuclear power stations could be built. (i) Suggest <b>two</b> arguments for building more nuclear power stations. [2] (ii) Suggest <b>two</b> reasons <b>against</b> building more nuclear power stations. [2]

The two key themes which have been identified across many of the subjects include the increase in the number of questions relating to real-world contexts, and the greater amount of choice available to candidates, both in terms of the different options within assessments and the methods by which they may display their skills.

Increasing use of real-world contexts can be illustrated from the study into **Mathematics**, where it was interesting to note that as early as 1957 one of the regulations sections stated that some of the questions might be set on the application of certain arithmetical processes to problems of everyday life in the home and the community. This appears to be one of the early explicit statements indicating a trend that became prevalent in testing all topic areas of mathematics in the GCSE Mathematics papers, although it was present even in the 19th century papers to some extent, especially in the area of Arithmetic.

An example of a question from the 1997 GCSE Mathematics assessment:

*Mrs McKenzie bought a large box of bags of crisps for her family. She told the children that the box should last 3 weeks if they ate 12 bags per week between them.*

*(i) How many weeks should the box last if the children eat 9 bags per week between them?*

*If the children eat  $n$  bags per week between them, the box will last  $W$  weeks.*

*(ii) Write down a formula which connects  $W$  and  $n$ .*

The studies investigated how topic areas within subjects have altered over the years. In this example from the **Geography** study (Figure 1), physical geography, human geography and geographical skills have featured since early days, but economic and environmental geography are more recent elements of the assessment.

In some instances practical considerations have affected the practice of assessment.



Photograph by Peter Asken. Cambridge Assessment Archives Ref: M/P 5/8

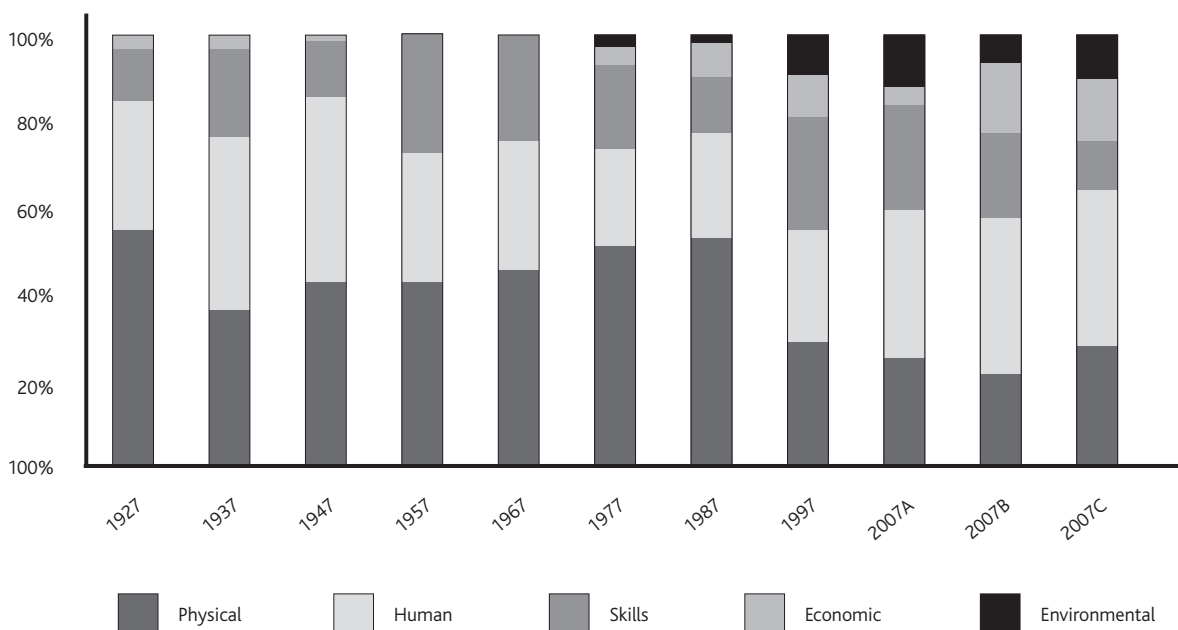
**Figure 2: Artwork in 1 Hills Road for marking**

For example, **Artwork** (Figure 2) used to be necessarily restricted by weight and size, because the work was sent to Cambridge and displayed in the Craft Hall at 1 Hills Road for marking.

*"Pieces of pottery must not exceed 12 ins. in any dimension, nor exceed 7 lbs. in weight. Pieces of sculpture or carving must not exceed three feet in any dimension nor exceed 20 lbs. in weight."*

1977 and 1987 Art specification

Now that schools themselves display candidates' work and examiners make visits to the schools, students' artwork is not limited in this way.



**Figure 1: Geography: Summary of topic areas over time**

**Table 2: Skills tested over time in English Literature**

<i>Skill</i>	1877	1887	1897	1907	1917	1927	1937	1947	1957	1967	1977	1987	1997	2007
grammatical analysis	✓		✓											
etymology	✓	✓	✓											
textual analysis	✓	✓	✓	✓										
scan (divide into metrics), knowledge of poetic/linguistic form (pentameter)		✓	✓	✓		✓	✓							
knowledge of author's life (external to text)			✓	✓		✓								
produce quotations verbatim	✓	✓	✓	✓	✓	✓	✓							
knowledge of literary, dramatic or poetic terms, concepts and mechanisms	✓	✓	✓	✓	✓	✓	✓	✓	✓					
translate text into contemporary prose retaining exact meaning	✓	✓				✓	✓		✓	✓				
comparison of text with factual information/external point of reference			✓	✓	✓	✓	✓				✓			
explain meaning of (extended) text (expound)	✓		✓	✓	✓	✓	✓		✓	✓			✓	
exact context of quote/excerpt	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓			✓	
overall evaluation of play/text/poem		✓	✓				✓	✓		✓	✓	✓		
give an account of a scene/sequence of events/story strand/poem	✓		✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
character analysis/development including comparison of characters	✓		✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
thematic analysis/overall theme					✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
qualities of the writing of the poem/text						✓	✓	✓	✓	✓	✓		✓	✓
views or attitudes of the author as detected from the text.						✓	✓	✓	✓		✓	✓	✓	
appreciation of text/describing impact upon self/reader response						✓	✓	✓	✓	✓	✓	✓	✓	
significance (importance) of a feature or scene of text							✓	✓				✓	✓	
understanding of staging of play/dramatic impact							✓		✓				✓	
reading comprehension of text (within 'extract based questions')											✓	✓	✓	✓
relating to characters											✓	✓		
imaginative writing – role playing a character in the text											✓	✓	✓	✓
unseen poem											✓	✓		

Finally, the studies have enabled analysis of the skills required by candidates at different points in time in specific subjects. In **English Literature** every question paper between 1877 and 1937 inclusive (Table 2) required candidates to quote verbatim from memory fairly substantial sections of the prescribed text. Earlier question papers used to require candidates to know the precise meaning, usage and etymology of words in the texts, and on occasion, questions would require candidates to quote a line in which a particular word appeared. Later question papers gave more emphasis to discussing overall meaning or themes of a text and describing or analysing the candidate's own response to a passage or character. A particularly common feature of later papers asked candidates to imaginatively play the role of a character in the text.

## Summary

The research proved a very interesting means of investigating the development of individual subjects. Naturally the method used –

sampling question papers from every tenth year- has some limitations. It is, for example, possible that short-lived topics or question paper structures have escaped our attention altogether. Also the researchers are unable to state for certain exactly when a particular change occurred – the research shows merely the first sampled year when such changes were seen.

However, many interesting details have emerged from every subject studied and two themes were repeated across many of the subjects. These were an increasing emphasis upon real-world contexts for questions in more recent years, and an increasing choice of topic areas and question/component options available to candidates.

For full reports in each of the seven subjects, please contact Gill Elliott, Assessment Research and Development Division, Cambridge Assessment, 1 Regent St, Cambridge, CB2 1GG. Email: [elliott.g@cambridgeassessment.org.uk](mailto:elliott.g@cambridgeassessment.org.uk).