

Do the questions from A and AS Level Economics exam papers elicit responses that reflect the intended construct?

Jackie Greatorex, Stuart Shaw, Phineas Hodson and Jo Ireland



CAMBRIDGE ASSESSMENT

Background

For a test to be useful a learner's achievement must be reflected in the test score. This achievement - the knowledge, skills and competencies that a test is designed to measure - is known as the construct. Questions on a test should measure the intended construct; the construct should be represented by the test. Tests can suffer from construct under-representation, where some of the relevant knowledge, skills and competencies of learners are not assessed.

Validity

Validity is the degree to which interpretations and uses of test scores are appropriate. For example, if there is strong evidence that A Level Economics is a good way of selecting students for university this interpretation/use has high validity. Validity is a vital quality of tests as it shows whether they provide useful information and can safely be used to make important decisions about people.

Validation

Validation is the study of whether a test has high or low validity for a use or interpretation. Many different types of evidence are needed to validate a test:

- questions can be studied to see whether they assess the construct;
- the accuracy with which grades are assigned to learners can be evaluated;
- decision making by selectors can be observed to show how evidence from a test is used.

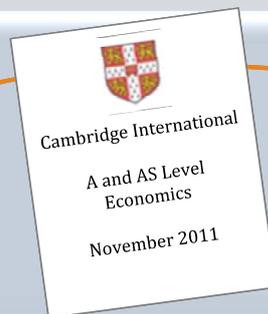
Finally, the level of validity can be judged by integrating these diverse pieces of evidence within a logical framework.

Method

Six experts in the qualification (Cambridge International AS/A Level Economics for 16-19 year olds) considered the Assessment Objectives from the syllabus, along with the questions from the examination.

They rated the extent to which each question elicited responses that reflected each Assessment Objective where 0 is 'not assessed at all' and 5 is 'strongly assessed'.

In total there were 71 questions and 2130 ratings.



Assessment Objectives

The Assessment Objectives (AOs) for the syllabus are:

AO1: Demonstrate knowledge and understanding of the specified content.

AO2: Interpret economic information presented in verbal, numerical or graphical form.

AO3: Explain and analyse economic issues and arguments, using relevant economic concepts, theories and information.

AO4: Evaluate economic information, arguments, proposals and policies, taking into consideration relevant information and theory, and distinguishing facts from hypothetical statements and value judgements.

AO5: Organise, present and communicate economic ideas and informed judgements in a clear, logical and appropriate form.

Results

Multiple choice questions

Question Papers 1 and 3

The syllabus states these questions will strongly elicit evidence relating to Assessment Objectives 1, 2 and 3

	Frequency of experts' ratings for multi-choice 0 = not assessed at all, 5 = strongly assessed					
	0	1	2	3	4	5
AO1	0	1	10	90	97	162
AO2	52	43	51	83	46	85
AO3	100	9	43	84	74	50
AO4	326	26	6	1	1	0
AO5	360	0	0	0	0	0

Most of the ratings for AO1 to AO3 (771 out of 1080) were 3 and above indicating the questions elicited the intended AOs. Most of the ratings for AO4 and AO5 (686 out of 720) were 0 indicating the questions did not elicit AO4 or AO5.

Data response questions

Question Papers 2 and 4

The syllabus states these questions will strongly elicit evidence relating to Assessment Objectives 2 and 3 and also Assessment Objectives 1, 4 and 5

	Frequency of experts' ratings for data response 0 = not assessed at all, 5 = strongly assessed					
	0	1	2	3	4	5
AO1	0	0	0	3	5	4
AO2	0	0	2	1	3	6
AO3	0	0	1	1	3	7
AO4	0	2	3	5	1	1
AO5	0	2	2	4	1	3

Most of the ratings for AO2 and AO3 (21 out of 24) were 3 and above indicating the questions elicited the intended AOs. Most of the ratings for AO1, AO4 and AO5 (14 out of 36) were 3 and above indicating the questions elicited these AOs to some extent.

Essay questions

Question Papers 2 and 4

The syllabus states these questions will strongly elicit evidence relating to Assessment Objectives 1, 3, 4 and 5 and also Assessment Objective 2

	Frequency of experts' ratings for essays 0 = not assessed at all, 5 = strongly assessed					
	0	1	2	3	4	5
AO1	0	0	1	10	10	33
AO2	38	4	4	2	6	0
AO3	0	0	0	4	16	34
AO4	5	8	10	11	15	5
AO5	0	1	3	5	16	29

Most of the ratings for AO1, AO3, AO4 and AO5 (188 out of 216) were 3 and above indicating the questions elicited the intended AOs. A few of the ratings for AO2 (8 out of 54) were 3 and above suggesting the questions elicited AO2 to a lesser extent.

Evidence for validity

The ratings reflect the weightings of AOs in the syllabus. At least one question was rated as eliciting each Assessment Objective. AO1 was more frequently assessed than any other AO, and AO4 was rated as relatively infrequently assessed. There may be legitimate reasons for differences in frequencies if certain skills are considered more important than others.

All three question types (multiple choice, data response and essay) elicited the construct(s) (AOs) described in the syllabus.

These findings contribute to providing teachers, students and higher education with evidence for validity of Cambridge International Examinations A and AS Level Economics.

Further reading

Crisp, V. & Shaw, S. D. (2011) 'Applying methods to evaluate construct validity in the context of A Level assessment', *Educational Studies*, Vol 38(2), pp. 209-222.

Greatorex, J., Shaw, S., Hodson, P. & Ireland, J. (2013) 'Using scales of cognitive demand in a validation study of Cambridge International A and AS level Economics', *Research Matters: A Cambridge Assessment Publication*, 15, 29-37.

Shaw, S. D. & Crisp, V. (2012) 'An approach to validation: Developing and applying an approach for the validation of general qualifications', *Research Matters: A Cambridge Assessment Publication (Special Issue 3)*.

Do the questions from A and AS Level Economics exam papers elicit responses that reflect the intended construct?

Jackie Greatorex, Stuart Shaw, Phineas Hodson and Jo Ireland



CAMBRIDGE ASSESSMENT

Background

At the heart of educational assessment lies an assumption that a test score indicates the degree of achievement (or attainment) in a particular content area. A test score should accurately represent a student's level of knowledge, skills or competencies that the test is designed to elicit – that is, the underlying constructs of interest. According to the *Standards for Educational and Psychological Testing*, a construct is “the concept or characteristic that a test is designed to measure” (AERA, APA, NCME, 1999, p.173). Thus it is important to establish that the test questions elicit performances that reflect the intended constructs. If the construct(s) is not well defined and the test questions are inappropriate, then it will be difficult to support the claims an Awarding Body wishes to make about the usefulness of its assessments, including claims that the tests do not suffer from factors such as construct under-representation (Messick, 1989). Construct under-representation indicates that the questions in the assessment fail to include important aspects of the construct. The credibility of educational examinations depends to a large extent upon a coherent understanding and articulation of the underlying latent abilities or construct(s) which they seek to represent.

Validity

Validity is the hallmark of quality for educational and psychological measurement. The claim of validity is that the examination adequately reflects the constructs and can be used as the basis for the inference of attainment or aptitude depending on the purpose of the test. The *Standards* (AERA et al., 1999) describe validity as “the most fundamental consideration in developing and evaluating tests” (p 9). Validity relates to whether the interpretations and uses of test scores are appropriate and meaningful (Messick, 1989; Kane, 2006). For this to be the case, various criteria must be achieved, such as good representation of intended constructs. Additionally, some conceptualisations of validity include consideration of the consequences that may result from the assessment, or at least the consequences that may result from the uses of assessment outcomes

Validation

A fundamental question within educational assessment relates to how (exactly) a claim to validity can be substantiated? Such questions are addressed through validation studies. Validation is “the process through which the validity of the proposed interpretation of test scores is investigated” (AERA et al., 1999, p.184). A validation study, such as the one described in part here, entails the collection of evidence in order to evaluate the proposed interpretation and uses of test scores. Clearly, “the provision of satisfactory evidence of validity is indisputably necessary for any serious test” (Hughes, Porter and Weir, 1988, p.4) in order to substantiate a claim of validity. Thus, ensuring that educational assessments have high validity is an essential aim of all those involved in the development of assessments.

A multi-method approach to validation, underpinned by validity theory, has been designed, piloted and implemented (Crisp and Shaw, 2011; Shaw and Crisp, 2012). The approach includes researching the question *Do the exam questions elicit responses that reflect the intended construct?* This study was based on the multi-method validation programme and was employed to investigate this research question which takes as its focus the Cambridge International Examinations A and AS Level Economics.

Method

Six experts in international Economics qualifications (for 16-19 year olds) considered the Assessment Objectives from the syllabus, along with the questions from the examination. They rated the extent to which each question elicited responses that reflected each Assessment Objective where 0 is ‘not assessed at all’ and 5 is ‘strongly assessed’. The method assumed that the assessment objectives were a representation of the underlying constructs. In total there were 71 questions and 2130 ratings. The ratings were in line with the statements in the syllabus.

Assessment Objectives

The Assessment Objectives (AOs) for the syllabus are:

Handout to accompany poster presented at BERA conference, 3 to 5 September 2013, University of Sussex

AO1: Demonstrate knowledge and understanding of the specified content.
 AO2: Interpret economic information presented in verbal, numerical or graphical form.
 AO3: Explain and analyse economic issues and arguments, using relevant economic concepts, theories and information.
 AO4: Evaluate economic information, arguments, proposals and policies, taking into consideration relevant information and theory, and distinguishing facts from hypothetical statements and value judgements.
 AO5: Organise, present and communicate economic ideas and informed judgements in a clear, logical and appropriate form.

Results

Multiple choice questions

Question papers 1 and 3

The syllabus states that these questions will particularly test Assessment Objectives 1, 2 and 3

Frequency of experts' ratings for multi-choice 0 = not assessed at all, 5 = strongly assessed						
	0	1	2	3	4	5
AO1	0	1	10	90	97	162
AO2	52	43	51	83	46	85
AO3	100	9	43	84	74	50
AO4	326	26	6	1	1	0
AO5	360	0	0	0	0	0

Most of the ratings for AO1 to AO3 (771 out of 1080) were 3 and above indicating the questions requested the intended AOs. Most of the ratings for AO4 and AO5 (686 out of 720) were 0 indicating the questions did not request AO4 or AO5.

Data response questions

Question papers 2 and 4

The syllabus states that these questions will particularly test Assessment Objectives 2 and 3 and also test Assessment Objectives 1, 4 and 5

Frequency of experts' ratings for data response 0 = not assessed at all, 5 = strongly assessed						
	0	1	2	3	4	5
AO1	0	0	0	3	5	4
AO2	0	0	2	1	3	6
AO3	0	0	1	1	3	7
AO4	0	2	3	5	1	1
AO5	0	2	2	4	1	3

Most of the ratings for AO2 and AO3 (21 out of 24) were 3 and above indicating the questions requested the intended AOs. Most of the ratings for AO1, AO4 and AO5 (14 out of 36) were 3 and above indicating the questions requested these AOs to some extent.

Essay questions

Question papers 2 and 4

The syllabus states that these questions will particularly test Assessment Objectives 1, 3, 4 and 5 and also test Assessment Objective 2

Frequency of experts' ratings for essays 0 = not assessed at all, 5 = strongly assessed						
	0	1	2	3	4	5
AO1	0	0	1	10	10	33
AO2	38	4	4	2	6	0
AO3	0	0	0	4	16	34
AO4	5	8	10	11	15	5
AO5	0	1	3	5	16	29

Most of the ratings for AO1, AO3, AO4 and AO5 (188 out of 216) were 3 and above indicating the questions requested the intended AOs. A few of the ratings for AO2 (8 out of 54) were 3 and above suggesting the questions requested AO2 to a lesser extent.

Evidence for validity

At least one question was rated as eliciting each Assessment Objective. AO1 was more frequently assessed than any other AO, and AO4 was rated as relatively infrequently assessed. There may be legitimate reasons for differences in frequencies if certain skills are considered more important than others.

All three questions types (multiple choice, data response and essay) elicited the construct(s) (AOs) described in the syllabus.

These findings contribute to providing teachers, students and higher education with evidence for validity of Cambridge International Examinations A and AS Level Economics.

References

- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (1999). *Standards for educational and psychological testing*. Washington, DC: American Psychological Association.
- Crisp, V. & Shaw, S. D. (2011) 'Applying methods to evaluate construct validity in the context of A Level assessment', *Educational Studies*, Vol 38(2), pp. 209-222.
- Greatorex, J., Shaw, S., Hodson, P. & Ireland, J. (2013) 'Using scales of cognitive demand in a validation study of Cambridge International A and AS level Economics', *Research Matters: A Cambridge Assessment Publication*, 15, 29-37.
- Hughes, A., D. Porter, & C.J. Weir. (1988). *Validating the ELTS test: a critical review*. Cambridge: The British Council and the University of Cambridge Local Examinations Syndicate.
- Kane, M.T. (2006). Validation. In R. L. Brennan (Ed). *Educational Measurement* (4th Edition) (pp.17-64). Washington, DC: American Council on Education/Praeger.
- Messick, S. (1989). Validity. In *Educational Measurement*, ed. R. Linn, pp.13-103. New York: Macmillan, pp.13-103
- Shaw, S. D. & Crisp, V. (2012) 'An approach to validation: Developing and applying an approach for the validation of general qualifications', *Research Matters: A Cambridge Assessment Publication (Special Issue 3)*.