



**Cambridge
Assessment**



**The Cambridge Approach
*to Assessment***



Principles for designing, administering and evaluating assessment

January 2009 – revised April 2017

Contents

- Section*
- 1 The Cambridge Approach**
 - 2 Cambridge Assessment, its principles and values**
 - 3 The range of assessments in Cambridge Assessment's work**
 - 4 The integrated model**
 - 5 Design and development of assessments**
 - 6 Validity**
 - 7 Reliability, precision and bias**
 - 8 Administration of assessments**

Reading this document

We have divided the content into sections to make it easy to read.

Sections 1 to 3 provide background to the Cambridge Approach.

Sections 4 to 8 give details of the principles which make up the Cambridge Approach.

Each section includes explanatory text, followed by clearly stated, numbered principles.

These principles form the key reference points for designing, administering and evaluating the assessments offered by the Cambridge Assessment Group.

The Cambridge Approach was drawn up by Tim Oates, Group Director of Assessment Research and Development, with the help and support of a wide range of professionals within the Group. Particular thanks go to: Mark Dowling, Elizabeth Gray, Nick Saville, Jean Marshall, Tom Bramley, Sylvia Green, Stuart Shaw, Di Palmer, Anthony Dawson, Hanan Khalifa Louhichi and Elizabeth Mills, for their reviews of the principles.

The Cambridge Approach was drawn up by Tim Oates CBE, Group Director of Assessment Research and Development. Any queries relating to this document should be addressed to him.

1 *Section 1*

The Cambridge Approach

Cambridge Assessment was established to promote educational excellence and high-quality learning through the use of assessment. In order to achieve this it is committed to ensuring that all its assessments are fair, have sound ethical underpinning, and operate according to the highest technical standards.

The Cambridge Approach sets out in a short and straightforward document what these principles involve in practice. It has been designed to be detailed enough to guide practice, but concise enough for the full content to be remembered, and to be fully read by busy people rather than placed on a shelf 'to be read at some point when I have enough time...'.

The principles emphasise strongly the need for manuals for individual assessments: detailed criteria and requirements linked to specific tests and examinations, with these reflecting their different purposes, measurement models, modes of assessment and modes of operation. These detailed criteria and requirements will change over time as measurement practice improves, as legal regulation changes, and solutions to operational problems are devised. They will be owned and maintained by the board within the Group responsible for each individual assessment.

The Cambridge Approach gives the guiding principles for assessment; the manuals embody these principles and provide the day-to-day reference for those designing and operating the organisation's tests and examinations.

Simon Lebus
Group Chief Executive

'The *Cambridge Approach* gives the guiding principles for assessment, set out in a short and straightforward document.'

2 Section 2

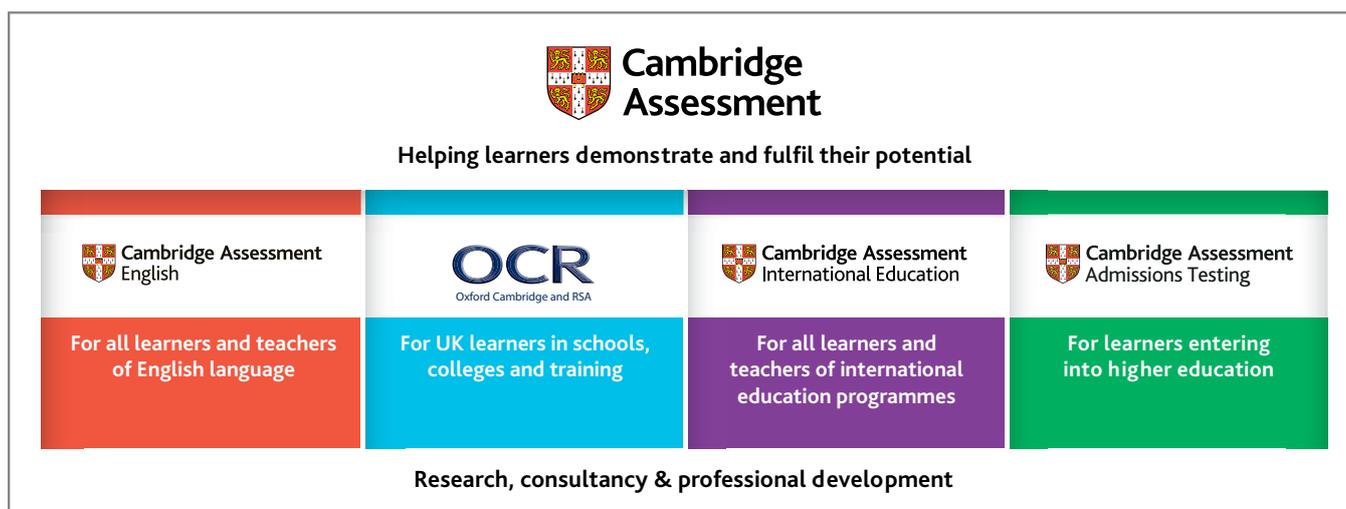
Cambridge Assessment, our principles and values

At Cambridge Assessment, our purpose is to help learners demonstrate and fulfil their potential. We care about making a difference for every learner.

As a department of the University of Cambridge, we provide education programmes and exams in over 170 countries offering global recognition. We unlock the power of education for millions of learners of all ages and abilities.

We have unrivalled depth of experience in national education systems, international education and English language learning. We are an international not-for-profit organisation with unique strengths and 160 years of expertise. Our qualifications are backed by the largest research capability of its kind.

We support and learn from teachers, schools and governments. Together, we are shaping education and creating a confident future for learners and a real and lasting impact on the world.



Organisational structure of the Cambridge Assessment Group

Our research underpins all our qualifications and education programmes. Across Cambridge Assessment we have a team of more than 100 researchers, which makes our research capability the largest of its kind. It is this research strength that enables us to help teachers, learners and governments stay at the forefront of education and unlock its power.

But our research is not just about ensuring our qualifications and services are the very best for learners. It's also designed to add to knowledge and understanding about

assessment in education, both nationally and internationally. We also carry out research for governments and agencies to inform their education reform programmes. It's all with one goal in mind – helping learners.

We strive to open doors for learners, to unlock the power of education and give them the confidence to thrive. We work with many national educational organisations and ministries through our international organisations, Cambridge English Language Assessment and Cambridge International Examinations. We work to improve standards of education, creating opportunity for learners around the world.

Cambridge Assessment has a high number of experts, with proven experience in curriculum and assessment design, and as a part of the University of Cambridge has access to world-leading resources, skills and research. We work in collaboration with institutions such as University of Cambridge Faculty of Education, Cambridge University Press and Fluentify to offer a comprehensive service to our partners.

In the UK we have been working with industry leaders to develop real industry projects for our Cambridge Technical and Cambridge National qualifications to give learners a head start in their chosen career.

3

Section 3

The range of assessments in Cambridge Assessment's work

Cambridge Assessment designs and administers assessments which have a formative function, as well as those with a predominantly summative function.

The assessment providers in the Group include:

Cambridge English Language Assessment

Tests and qualifications from Cambridge English Language Assessment are taken by over five million people, in 130 countries each year. Cambridge English's teaching qualifications provide a route into the English Language Teaching profession for new teachers as well as career development opportunities for experienced teachers. Cambridge English's qualifications and development work are underpinned by an extensive programme of research and evaluation work, undertaken by a dedicated research unit.

Cambridge International Examinations

Cambridge International Examinations is the world's largest provider of international qualifications for 5 to 19 year olds. Cambridge qualifications are available in over 160 countries.

Oxford Cambridge and RSA

Oxford Cambridge and RSA (OCR) is a leading UK awarding body providing general academic and vocational qualifications for young people and adults.

Assessment Research and Development

The Assessment Research and Development division supports development and evaluation work across the Cambridge Assessment Group.

The providers in the Group develop, administer and evaluate a very wide range of assessment instruments, from essay-based papers for general education to performance assessment in occupational settings.

Assessment takes place in the context of national and international legislation, and assessment/qualifications codes and criteria – the latter being predominantly national in character. Assessment development and administration should take due account of legislation and pending changes in legislation. Assessment/qualifications codes and criteria should be viewed as necessary operating frameworks, themselves undergoing change as knowledge of assessment and public policy improves. The development and evaluation of assessment by Cambridge Assessment itself contributes to the corpus of knowledge on assessment. This can yield valuable evidence for the refinement of codes and criteria, as well as new approaches not yet anticipated by, or incorporated in, these codes and criteria – for example, use of item types not yet used in a given subject or occupational area, new types of grading models, etc.

Care should thus be taken to comply with the requirements of codes and criteria, but it is essential to collect evidence from assessment development, administration and evaluation which can be used, by those responsible for the codes, to refine them.

The principles outlined in this document apply to the full range of assessment instruments developed, administered and evaluated by the assessment providers in the Group. These fall into the following main categories:

Objective-response items

Where candidates are required to provide a highly defined response, such as adding a line to a diagram, completing a number sequence, completing a statement.

Multiple-choice items

A specific form of objective-response item, where the candidates' responses are constrained to a given list of alternatives, selecting the response which best completes a statement or answers a question.

Short-answer items

Where candidates have to provide succinct, specific responses to questions, but where the marking scheme allows some variation in those responses.

Extended-response items

Where candidates have to create a given amount of text and/or working, and which allow candidates to organise their responses and present material in an original way.

Coursework/non-examined and teacher assessment

Where candidates produce written reports or complete structured activities which are typically marked by teachers and tutors who are also responsible for delivering the learning programmes in which the assessment is located.

Performance assessment

Where performance in work tasks, or activities such as laboratory work are observed and scored/rated by a teacher or qualified assessor. Assessment in work contexts typically is guided by occupational standards.

Evidence accumulation

Where candidates are responsible (with appropriate levels of support) for gathering evidence of performance in a range of activities and for assembling this evidence into a portfolio/record of evidence. This is then scored/rated by teachers or qualified assessors and subject to external moderation/verification.

Across these, there are three typical development and management models:

1 Banked items

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

This can be characterised as a 'data accumulation' model. A key aspect of such an approach is linking items to specific skills, knowledge and/or understanding and placing items on specific measurement scales associated with these constructs. Items typically are subjected to pre-testing in order to yield item performance data, allowing inclusion in the bank once specified quality criteria have been met. The administration of the items tends not to include an awarding session, and (within secure systems), repeated administration of the item yields more data which are 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. Overexposure and material becoming out of date are key threats to validity in this model.

2 Awarding-based assessments

design – review – administration – outcomes – awarding – endorsement

This can be characterised as a 'review and award' model. Items typically are not subjected to pre-testing, and the total score in an examination or examination component is scrutinised in awarding processes, principally 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. Endorsement of the award is an important post-award process. The main evaluation processes comprise separate review and dedicated processes such as comparability exercises. Routine reuse of items is not a feature of this model. Without pre-testing, unanticipated problems in the characteristics of the item (bias, wrong level of difficulty) are key threats to validity in this model (see section 6).

3 Performance-based assessments

specification of standards/objectives – identification or specification of opportunities to collect evidence – judgement – moderation

This can be characterised as an 'evaluation' model. Assessment is based on simulated or naturally occurring activities which allow assessment against stated standards or assessment objectives. Samples of evidence or assessment decisions can be subjected to external moderation. Overall evaluation is discharged through separate review – this review can include data-oriented evaluation where sufficient numbers of assessments have been made to enable a scrutiny of possible gender bias etc. and where the assessment is expected to correlate with later performance. The standards/objectives can be pitched at very different levels of detail (e.g. highly specific requirements such as making an accurate medical diagnosis of a specific condition through to very generic requirements such as 'exercising customer care'). The emphasis on performance in naturally occurring activities or simulations approximating to naturally occurring activities tends to involve relatively open and complex assessment tasks/opportunities in which assessors have to make judgements based on the stated standards. Variability in the interpretation of the standards and the openness/complexity of the elements in the tasks are key threats to validity in this model.

The principles in this document apply to all modes of assessment – paper, on-screen, adaptive, by observation, oral questioning, and so on, and all 'item types'.

4 Section 4 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 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 Documentation 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
-

***Integrated model:
Principle 1***

At all stages of this integrated process, comprehensive documentation should be accumulated, by nominated officers, on the characteristics of the assessments and their constituent elements. The documentation should be structured using the framework 1–10 (see above), and be informed by the principles which are presented in the following sections of this document. This will constitute the manual for each assessment or group of related assessments. The documentation should be sufficiently detailed and precise to support effective audit processes and accountability arrangements.

The manual for each assessment is a vital tool for effective development, administration, evaluation and refinement. It supports:

- audit and other review and accountability processes
- review of allied and cognate assessments
- assessment management and administration processes
- historical review of the 'evolution' of an assessment

For some assessments and qualifications, the manual will be a consolidated document. In other cases the data and documents may be dispersed. In such circumstances the 'manual' will comprise of a succinct index of the data and documents, and all elements will be reviewed to ensure that they meet the criteria presented in this document. Records and documents such as those developed to meet quality criteria and standards such as those from ISO are intended to be a part of this and, where they meet the requirements presented in this document, do not require duplication in a different form.

For all phases of assessment design, operation and evaluation, information on proposed and actual candidates is vital, and should form a part of the documentation associated with the qualification.

***Integrated model:
Principle 2***

Data should be accumulated on the characteristics of the candidates taking the assessments. This is essential for informing issues of standards, bias, equity etc. This becomes particularly important when qualifications are in pilot phase, and when change in specifications is implemented.

5 *Section 5*

Design and development of assessments

The key criterion driving assessment at Cambridge Assessment is validity. Design of specific instruments is determined by: purpose; the characteristics of the setting in which the assessment will operate; and any national codes and criteria which specifically apply.

The Group operates across a very wide range of national contexts, and the design of assessments and their administration arrangements must be sensitive to the specific conditions which apply in each of those contexts.

The principles laid down below address different aspects of validity but should be seen as different elements of a 'thread' of validity which runs continuously from the specified purpose of an assessment, through the specification for the assessment, the items which make it up, the inferences which are drawn from it, and its impact.

The Cambridge Approach emphasises the importance of viewing 'fitness for purpose' as being related to the interests and needs of the various stakeholders in the assessment and not merely some of them. Assessments and qualifications which are used for selection and progression purposes – e.g. for university admissions – must have measurement properties consistent with the use to which they are being put. But it is vital also to recognise that formative assessment and in-course assessment, while not used for selection purposes, can have a powerful effect on learners' views of themselves (their identity as learners, their sense of belonging to specific learner groups) and can dramatically affect learners' decisions about the things at which they are 'good' and 'bad'. This can affect learners' decisions about which directions to take in learning programmes, where to put effort in order to excel etc.

As a result, Cambridge Assessment does not differentiate different levels of quality assurance, development effort and evaluation/research effort on the basis of different 'classes' of assessment. It believes that misclassification and lack of validity are serious matters in all assessments, and effort should be devoted to (i) maximisation of validity, and (ii) quantification and reduction of measurement error and bias, in all assessments.

Design: Principle 1

Design processes for assessments from Cambridge Assessment must be predicated on:

- 1 Identification of clear need for the assessment, to the benefit of the individual, society and the economy
 - 2 Precision in the specification of the purpose, content, mode, administration processes and results service of the assessment. The principles on validity included in this document emphasise the importance of matching mode of assessment, item type and content, rubrics and regulations etc. to the purpose of the assessment and the context in which it is intended to be administered. Precision in specifying these aspects of the assessment during the design process enhances its integrity and robustness. The expectations regarding measurement quality and necessary characteristics of the assessment are key aspects of the 'design blueprint' for the assessment.
-

-
- 3 Consultation with all relevant parties

 - 4 A development process which has sufficient resources, has an adequate timeframe and has clear criteria for the evaluation of the adequacy of the emerging assessment. The specification for each assessment will be made available in its final form with sufficient lead time to allow those administering the assessment to make adequate preparation for it – for example, to put well-designed learning programmes in place.

 - 5 A design phase which guards against unnecessary divergence of approach across the Group, by comparison of the emerging assessment with any like developments in the Group, adoption of all state-of-the-art approaches, and clear statement of fitness for purpose.
-

Cambridge Assessment is committed to development and piloting processes which are designed to enhance an assessment at all stages of its development:

Design: Principle 2

Concurrent formative evaluation should be used, wherever appropriate, to provide focused support to the development process.

Design: Principle 3

All manuals for assessments should specify initial, medium and long-term validation and evaluation strategy for the assessment.

Design: Principle 4

Care should be taken to comply with the requirements of national and international codes and regulations, criteria and frameworks; but it is essential to collect evidence from assessment development, administration and evaluation which can be used, by those responsible for the codes, to refine them.

Design: Principle 5

Initial cycles of the operation of an assessment should be subject to adequate evaluation. Evaluation information from the operation of the initial cycles of the assessment should be fed into validation and refinement of the qualification.

Design: Principle 6

Data from the assessment and data from other sources should be used to ensure the most precise initial fix of standards in the first assessment cycles.

Design: Principle 7

All evaluation work should be informed by the fullest information on the candidature for the assessments, recognising that for some assessments, and in some country contexts, systems for the provision of this kind of data are underdeveloped.

Design: Principle 8

Ethical safeguards should be put in place for all learners taking part in piloting processes for new assessments, including protections (dispensations and compensations) which will safeguard candidates in the event of problems arising during piloting.

Design: Principle 9

Regulations relating to security should be observed throughout all initial and recurrent development work – such as continuing development of assessment items. Items and associated materials should be kept secure at all stages of development. Conflicts of interest for those involved in the development of assessment items should be avoided.

Design: Principle 10

The presentation of assessments will conform with house style, and with international best practice in respect of such matters as scientific units and presentation of diagrams. The assessments will be free from content errors and typographical errors. Assessments will be designed with sensitivity to the gender, nationality, and ethnic and religious backgrounds of those taking the assessments.

6 *Section 6* Validity

Cambridge Assessment treats validity – and validation processes – as a pervasive concern which runs through all of its work on the design and operation of assessment systems. This begins with consideration of the extent to which the assessment is assessing what it is intended to assess and flows out to the uses to which the information from the assessment is being put. The extent to which the inferences which are made on the basis of the outcomes of the assessment are meaningful, useful and appropriate is seen as a vital aspect of validity. Assessment is located in context, and validity is linked to the use of a specific assessment in a specific context, with a particular group of candidates. Validation consists of the collection and analysis of data which address the validity of the assessment and, in particular, whether the inferences made on the basis of the outcomes of the assessment are appropriate.

This approach yields a number of principles:

Validity: Principle 1

The manual for the assessment and any supporting material for the assessment should include a clear statement of purpose – what the assessment is intended to assess and the uses to which the information from the assessment should be put, including arguments against any known misuses or potential obvious misuses. While Cambridge Assessment cannot be held directly responsible for misuses of the information from the assessments for which it is responsible, the approach to validity adopted by the organisation emphasises that all reasonable steps should be taken to ensure that the information from the assessment is used appropriately. Suitable information and guidance should be provided for all assessments regarding use of outcomes; adequate contractual safeguards should be put in place where appropriate; and use and impact studies routinely commissioned. Impact studies should include investigation of any effects on teaching and learning, washback into the curriculum, etc. using the stated purpose of the assessment as a reference point. Assessment developers and administrators should be aware that the uses to which the outcomes of assessments are put are within the definition of validity used by Cambridge Assessment.

Validity: Principle 2

Each assessment should provide a basis for making valid inferences about the skills, knowledge and/or understanding which forms the focus of the assessment – the specific constructs at the heart of each assessment. The assessment should possess well-founded content validity. This entails using existing and/or original research to establish what elements of knowledge, understanding and/or skills are the intended construct. Understanding of what makes up effective performance in physics, languages etc. is constantly changing; as is the content of educational programmes, societal expectations and requirements in the economy. Those responsible for developing assessments should ensure that the constructs at the heart of each assessment are well grounded – that they are indeed the things which make up effective performance in the subject or occupational area. The constructs at the heart of each assessment should be clearly stated in the manual for the assessment. This should include a reference to any key research or sources which were used as the rationale for including the constructs in the assessment and for determining the way that they are framed

Validity: Principle 3

The assessment should sample adequately from a subject area or area of performance in order to make claims about processes in such domains as well as content. Some assessments may require recall of facts while others require a focus on application and analysis or on procedures and cognitive processes. Where appropriate, the assessment should require sufficient processing of this kind to ensure that a reasonable claim can be made that the candidate has demonstrated use of such processes. While sampling from domains should be at an adequate level, assessments with more than one component should avoid overassessment by seeking to avoid unnecessary repetition of assessment of the same knowledge, understanding and skills.

Validity: Principle 4

The internal structure of the assessment should be consistent with the internal structure of the content domain. Those things which attract marks or credit should be genuinely significant elements of performance in a subject, and the assessment should have authentic sequencing of activities and processes.

Validity: Principle 5

'Generalisability' should be addressed in the development and evaluation of the assessment. It should have sufficiently representative coverage of the content and processes of the construct domain to allow reasonable claims regarding generalisability – the extent to which the skills, knowledge and other attributes can be reproduced in settings other than the setting of the assessment.

Validity: Principle 6

'Construct under-representation' should be guarded against in the development and evaluation of the assessment. This occurs when the tasks which are measured in the assessment fail to include important dimensions or facets of the construct which is the focus of the assessment. Under such circumstances, the test results are unlikely to reveal a candidate's abilities in respect of the construct.

Validity: Principle 7

'Construct-irrelevant variance' should be minimised. For example, if an examination question is about ratio, it should only be possible to score on the item by using ratio. Any extraneous elements – such as obscure contexts which are used in the question, or asking the question to be completed with undue speed – should be identified and remedied. Too high a reading level demanded by the question and/or the stimulus material, undue complexity of stimulus material, and poor question layout are

common problems which should be guarded against. It is essential that any increase in attainment/performance should be directly linked to elements of the constructs which are the focus of assessment, and not derive from construct-irrelevant difficulty. 'Construct-irrelevant easiness' should be avoided, which occurs when extraneous clues in item or task formats permit some individuals to respond correctly or appropriately in ways that are irrelevant to the construct being assessed. 'Construct-irrelevant difficulty' should also be avoided, and occurs when extraneous aspects of the task make the task in an irrelevant manner more difficult for some individuals or groups. While the first type of construct-irrelevant variance is associated with a score higher than one would attain under normal circumstances, the latter is associated with a notably lower score.

Validity: Principle 8

The measurement model, assessment approaches, grading and reporting arrangements should be well matched to the constructs at the heart of the assessment and the uses to which the information from the assessment is to be put. The measurement characteristics of the assessment should be established, using techniques matched to the purpose and form of the assessment – for many assessments, investigating the psychometric properties of the assessment is an essential part of this.

Validity: Principle 9

Validation studies should be commissioned, where necessary, to ensure that the intended inferences can be made from the assessment, and that the assessment is performing as expected. For example, if the assessment is intended to predict later performance in a following stage of education or training, the relationship between the outcomes of the assessment and the outcomes of the subsequent education and training programme should be systematically explored. Studies should establish the extent to which the scores derived from the assessment are in reasonable accordance with scores, behaviours etc. from other assessments or other settings, which relate to the constructs which are the focus of the assessment.

Validity: Principle 10

The different development teams in Cambridge Assessment should endeavour to make links across the organisation to ensure that cognate assessments are treated in a suitably similar way, and that lessons learned from the development of different assessment instruments for similar constructs are used to inform all related assessments.

7 *Section 7* **Reliability, precision and bias**

Cambridge Assessment links validity and reliability closely together. While validity relates to whether the assessment is assessing what it is intended to assess and whether the inferences drawn from the results are well founded, reliability relates to the stability of the assessment, for example, whether on different occasions or using different markers the same outcomes are rated in the same way. If validity is poor, reliability in the assessment will be to little effect. If reliability is poor – and results are unstable – validity is compromised.

They are different, but equally important aspects of assessment.

The reliability of an assessment relates to its stability, consistency and precision – how it can provide repeatable outcomes for candidates with comparable characteristics (knowledge, skill, understanding) at different times and/or places.

Reliability: Principle 1

Developers, operations managers and evaluators should seek to understand the limitations of precision in an assessment, employing appropriate techniques to quantify and express clearly the estimates for reliability of assessments, with due attention to the purpose of the assessment and its uses. This should be provided by analysis at item level as well as any aggregated level of reporting.

Reliability: Principle 2

Developers, operations managers and evaluators should seek to set the precision of an assessment at a level consistent with its purpose.

Reliability: Principle 3

Developers, operations managers and evaluators should fully document the approach to estimating reliability and make a clear report of the limitations on precision to users of the outcomes of the assessment.

Discrimination is a fundamental purpose of assessment – distinguishing those who are competent at x from those who are not, those who can demonstrate knowledge of x and those who cannot, etc. This is fundamental to assessment – both summative and formative in purpose. However, alongside discrimination can appear bias – unwanted effects in assessment which give rise to unfairness. Bias occurs where there are systematic differences in assessment outcomes which are associated with being a member of a specific group – evident when people from two groups are in the same position in respect of the constructs which are the focus of the assessment, but have been given different scores. The existence of differences between groups is not unequivocal evidence of bias, since there are many ways in which members of different groups do, in reality, vary. Conversely, the lack of difference does not prove definitively that an assessment is not biased. Sensitive analysis for bias is thus a vital part of test development and evaluation, and appropriate action to remedy bias should follow.

Reliability: Principle 4

Developers, operations managers and evaluators should ensure that bias is monitored through the collection of adequate data and sufficiently analytic examination of those data.

Reliability: Principle 5

Operations managers should ensure that action is taken to remedy bias and the effects of bias where it is detected.

Concern for the detection and remediation of bias should not be limited to the technical characteristics of assessments. Bias can enter into assessment systems in respect of access and opportunity prior to the test – for example, entry decisions and approaches to preparation for the assessment. While Cambridge Assessment staff cannot take direct responsibility for remedying the causes of bias in this extended sense, it is important to recognise that this kind of bias can occur, that analysis of the outcomes of assessments can frequently help with the detection of this kind of bias, and that the results of such analysis should be made available to those who can take direct responsibility for remedying such bias.

8

Section 8

Administration of assessments

The manual for each assessment is the key reference for the administration of the assessment. The principles laid out below are critical in respect of the full range of assessments administered by the Group – ranging across all classes of paper-based and electronically administered assessments.

***Administration:
Principle 1***

The assessment should be administered in line with the procedures stated in its manual, including all security arrangements, the regulations for administration of special arrangements and the operation of special considerations.

***Administration:
Principle 2***

All assessments will be accompanied by adequate supporting materials in media readily accessible to those administering the assessment and to those taking it. This should include a clear statement of the purpose of the assessment; clear guidance on its availability and pattern of administration; sufficient advance notice of withdrawal of the assessment or its replacement; and other key elements of the regulations for the assessment. Specimen and exemplary materials should offer clear, helpful guidance to those administering the assessment and those taking it, and should be well matched to 'live' materials.

***Administration:
Principle 3***

Assessment administration should take place in circumstances and an environment which do not impinge adversely on candidates' capacity to perform in the assessment; variation in circumstances and environment should be managed such that they are not a threat to validity.

***Administration:
Principle 4***

All security requirements applying to the assessment should be observed. Contingency plans should be in place and documented in the manual for the assessment, in order to allow swift and effective response to security breaches. Potential risks to security should routinely be reviewed, and revised procedures promptly put in place and included in the manual for the assessment.

***Administration:
Principle 5***

Clear instructions on how to make responses should be provided, with practice materials and practice sessions available where this provision is consistent with the aims of the assessment.

***Administration:
Principle 6***

Where irregularities or disruptions to test administration occur, procedures should be in place for: evaluating the scale and consequences of the irregularity or disruption; activating corrective or compensatory actions; recording all details of events.

Administration:
Principle 7 Incident procedures should be implemented in line with the procedures stated in the manual for the assessment, with all insights from the incident used to enhance incident management procedures. Where revised arrangements become necessary to reduce risk, manuals/guidelines should be promptly updated.

Administration:
Principle 8 Robust quality control and quality assurance processes should be in place in respect of marking procedures. Marking and scoring errors should be monitored and their frequency recorded. This information should be used as part of action to minimise errors.

Administration:
Principle 9 The manual for the assessment should include clear regulations for dealing with loss, absence or late arrival of evidence from assessments.

Administration:
Principle 10 Results should be published in accordance with individuals' rights to confidentiality and to the schedule stated in the manual for the assessment.

Administration:
Principle 11 Where information from assessments is made available to teachers and trainers for purposes such as reflecting on the effectiveness of learning programmes, this information will be made available in the form and to the schedule outlined in the manual for the assessment. Clear guidance will be supplied which states the legitimate uses to which the information can be put, with a strong emphasis on valid inferences and use. The utility of such information will be regularly reviewed, and guidance and regulations promptly updated in the light of the findings of the review.

Administration:
Principle 12 Review and appeals procedures should be stated clearly in media to which candidates and those administering assessments have ready access. The review and appeals procedures should be conducted in accordance with the regulations laid down in the manual for the assessment.

Administration:
Principle 13 The candidates' original outcomes from the assessment will be retained for reference until results are issued and for a sufficient period thereafter to allow for review and appeals procedures to be completed. Some outcomes may be retained for research and other purposes. Suitable consents will be obtained for the uses to which the data are put.

Administration:
Principle 14 Where assessment materials are subject to high security and problems of exposure (e.g. banked items) any necessary external review should be undertaken by third parties nominated for their expertise and their capacity to preserve the security of the assessment materials.

Conclusion

The principles contained in this document collectively form the Cambridge Approach and, in line with the commitments to continuously improve assessment, will themselves be subject to review and revision. By contributing to the quality of the design, administration and evaluation of assessment, the Cambridge Approach will help to assure that the assessments provided by the Group continue to meet the needs of individuals, society and the labour market, nationally and internationally.

Every year eight million learners take our exams in 170 countries, and with 160 years' experience and the largest research facility of its kind in Europe, the Cambridge Assessment Group is chosen to be a key education partner to governments across the globe.

We are committed to education excellence and are trusted by universities, colleges, employers and governments, and are proud of the relationships we build and partnerships we form around the world.

At Cambridge Assessment, the reliability of our assessments stems from evidence-based and research-led approaches to all products, services and new developments. We have the largest research capability of its kind in Europe with more than 110 researchers across the Group.

Our researchers conduct and publish authoritative research in order to validate, improve and develop our assessments and services, and to influence thinking and policy on educational assessment nationally and internationally.

cambridgeassessment.org.uk
@Cam_Assessment



UNIVERSITY of CAMBRIDGE
Local Examinations Syndicate

Cambridge Assessment is the brand name of the University of Cambridge
Local Examinations Syndicate, a department of the University of Cambridge.
Cambridge Assessment is a not-for-profit organisation.

© UCLES 2018

