

Cambridge's approach to Assessment

Principles and practicalities: an overview of
principles, marking and grading

Anne Gutch
Acting Director, Assessment

July 2018



Contents

- ▶ Underlying principles of assessment procedures
- ▶ Stages in the assessment process
 - Devising/ updating syllabus
 - Scheme of assessment
 - Producing Assessment material
 - Administration of the exam/test
 - Reliable marking and moderating
 - Grading – setting and maintaining the standard
- ▶ Primary and Lower Secondary programme update
- ▶ Any questions?

Producing fair assessments

- ▶ For tests to be **fair and reliable**, we need to get four assessment processes right:
 - ▶ Syllabus
 - ▶ Question paper / test
 - ▶ Marking
 - ▶ Grading

Activity

- ▶ Discuss what types of test format would and would NOT be an appropriate option for assessing: consider multiple methods

- ▶ Art and Design
- ▶ Music
- ▶ English
- ▶ French
- ▶ Maths
- ▶ Chemistry

Activity (cont.)

Here are some ways of assessing: consider

Art

Music

English

French

Maths

Chemistry



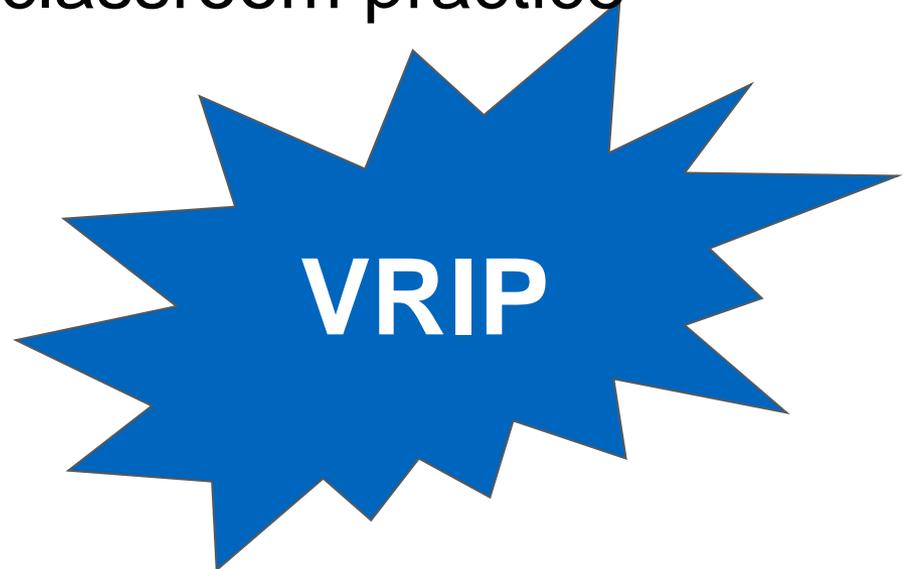
Coursework



Fair Assessments

We strive for:

- ▶ ...**v**alidity
- ▶ ...**r**eliability
- ▶ ...positive **i**mpact on classroom practice
- ▶ ...**p**racticality



Devising and updating Syllabuses

Considerations:

- ▶ Candidates: what are their needs and backgrounds?
- ▶ Is the syllabus fit for purpose?
- ▶ Does the syllabus state what it is setting out to test?
- ▶ Are the tasks fit for purpose for the type of assessment? Are they **valid** and do they have positive educational **impact**?
- ▶ Should it include coursework?
- ▶ Does the syllabus content **sample** the subject appropriately for the level? (Content validity)
- ▶ Is the content **up-to-date**? Does it reflect latest thinking?
- ▶ Does the content permit **progression** to a higher level?
- ▶ Does the content reflect Cambridge Learner attributes?

Coursework in summative assessments

Activity

- ▶ Two perspectives:
 - ▶ Teaching and learning
 - ▶ Assessment

- ▶ Consider: pros and cons of the use of coursework
 - ▶ for learning and teaching
 - ▶ for use in Assessment

Test Specification/Scheme of assessment

Is the scheme of assessment going to produce assessment that is **valid, reliable and practicable** for Centres, and likely to have a good **impact** on teaching?

- ▶ Some considerations:
 - ▶ length of exam
 - ▶ number and type of components
 - ▶ number and type of items
 - ▶ manageability in large schools in different settings
 - ▶ weighting of components in assessment
 - ▶ examiners can mark it reliably.

Producing Assessment material

Must meet requirements regarding:

- ▶ Validity: conforms to syllabus
- ▶ Validity: avoids irrelevant matters, testing what it is meant to test
- ▶ Validity: avoids construct under-representation (the test is too narrow in focus)
- ▶ Coverage: covers a fair proportion of the syllabus
- ▶ If there is a choice of questions, they are of similar demand
- ▶ Discrimination: (distinguishes between different abilities)
- ▶ Appropriate level overall but allows performance at all levels (not too easy or too hard)

Producing Assessment material (cont.)

Must meet requirements regarding:

- ▶ Use of accessible and clear language
- ▶ Factual correctness and accuracy
- ▶ Consideration of impact on students e.g. topics
- ▶ Cultural appropriateness for wide international candidature or where country specific
- ▶ Guidelines needed to support writers



Producing Assessment material (cont.)

Mark Schemes

Developed alongside the question papers; part of assessment instrument

As with question papers – must be:

- ▶ Accurate
- ▶ Lead to consistent marking
- ▶ Conducive to discrimination
- ▶ Complete and comprehensive
- ▶ Clear for markers

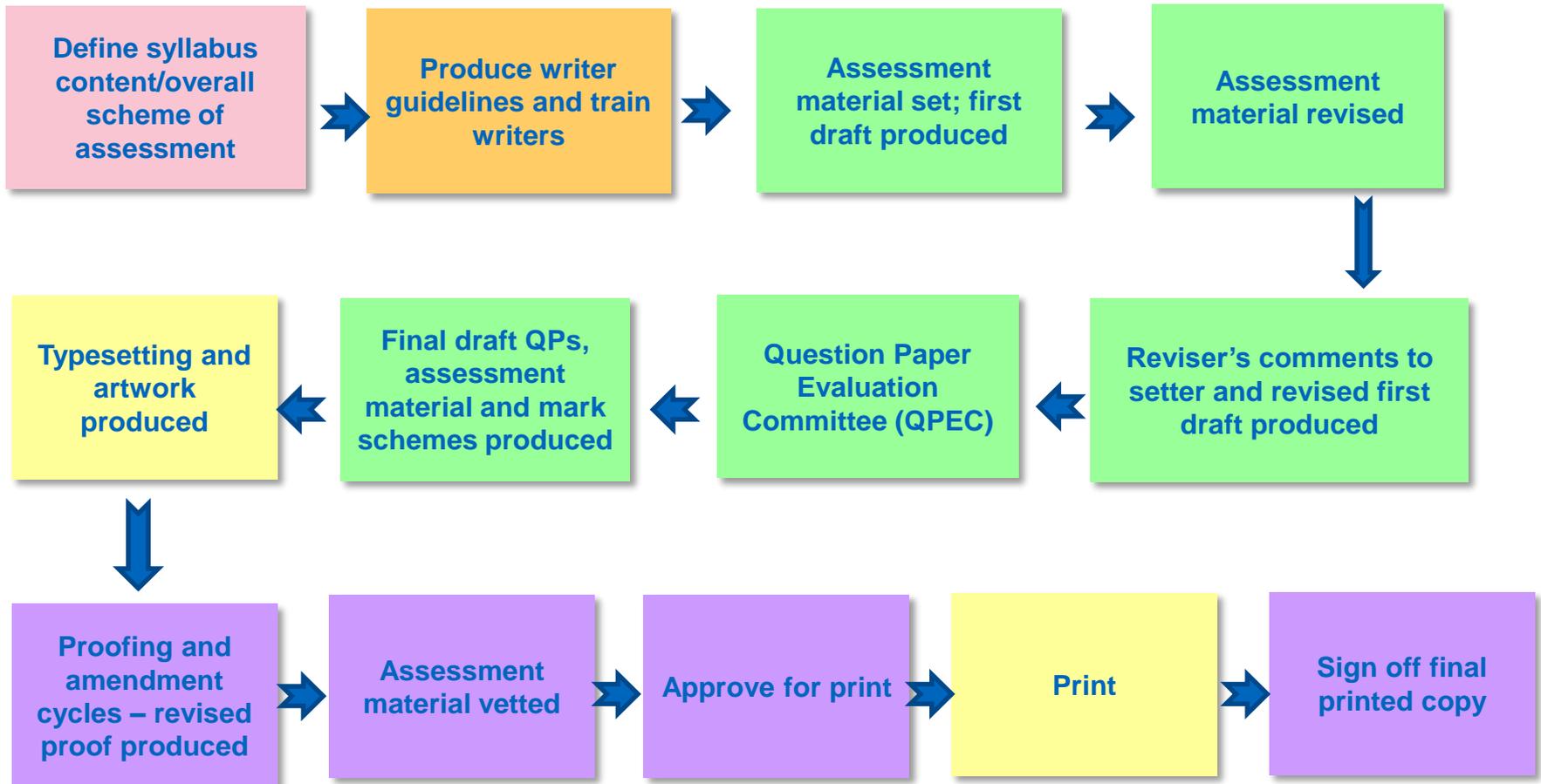
Producing Assessment materials (cont.)

Quality Assurance of process

- ▶ Proscribed and documented process
- ▶ Multi stage
- ▶ Defined personnel e.g. content versus typographical checking
- ▶ Professional typesetters and artists
- ▶ Training of participants

Producing Assessment materials (cont.)

Process



Administration of the exam/test

- ▶ Production of Guidelines / Handbook
- ▶ Timetabling: setting an international timetable, windows and fixed dates
- ▶ Suitable test locations and conditions
- ▶ Standardised conditions
- ▶ Security: training and monitoring of invigilators; secure storage at schools; secure despatch / transfer of materials
- ▶ Special consideration arrangements

Reliable marking and moderating

Types of marking:

- ▶ On a cline: objective to subjective
 - ▶ Multiple choice
 - ▶ Short answers e.g. Science, Maths, Languages
 - ▶ Longer answers e.g. History
 - ▶ Essays e.g. Literature

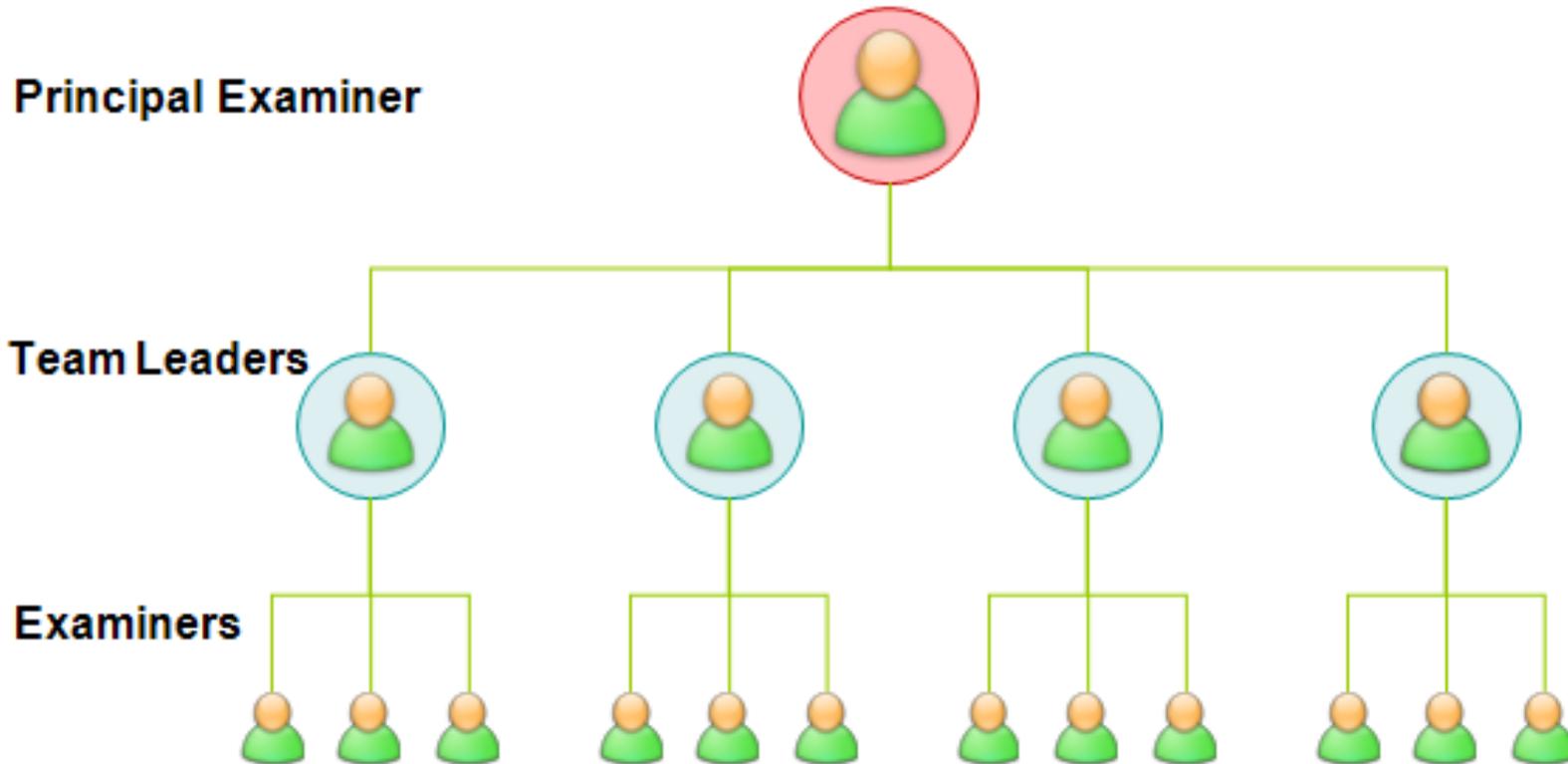
AND

- ▶ Oral responses
- ▶ Music performance
- ▶ Art

Reliable marking and moderating (cont.)

- ▶ Comprehensive mark scheme so all examiners award marks in the same way
- ▶ On line marking
- ▶ Appropriate number of scripts allocated to examiners
- ▶ Standardisation Meetings/process to check common understanding/markings to same standard
- ▶ Marking of common practice scripts before marking starts
- ▶ Feedback on initial marking
- ▶ Arithmetic checking
- ▶ Moderation of teacher-marked components
- ▶ Grade Review: re-marking of sub-set of 'borderline' candidates

Reliable marking and moderating (cont.)



Grading

- ▶ Means deciding on grade thresholds- the minimum number of marks needed to award the grade
- ▶ Issues to consider:
 - ▶ maintenance of **standard** over time
 - ▶ comparability of thresholds across successive test sessions
- ▶ The awarding ***standard*** is the level of performance which a candidate must demonstrate in order to be awarded a particular grade.

Why is grading needed?



Rohit got **124 marks** in his A Level Physics exam in **June 2017**

Amit got **129 marks** in his A Level Physics exam in **June 2016**

Who is better at Physics?

Why is grading needed??



Rohit got a **grade B** in his A Level Physics exam in June 2017

Amit got a **grade C** in his A Level Physics exam in June 2016

Who is better at Physics?

The two key questions in grading

- 1. How difficult is this year's exam compared to last year's?**
- 2. How good are this year's candidates compared to last year's?**

How difficult is the exam?

What evidence do we consider?

Examiner judgement

Item-level data

What do we do with the answer:

If this year's exam is the same difficulty as last year's?

If this year's exam is more difficult than last year's?

If this year's exam is easier than last year's?

How strong are the candidates?

What evidence do we consider?

Forecast grades

Entry patterns

Prior attainment

Performance on other papers

What do we do with the answer:

If this year's candidates are the same ability as last year's?

If this year's candidates are stronger than last year's?

If this year's candidates are weaker than last year's?

Activity: Grading an exam

- ▶ The mean mark is lower than last year.
- ▶ The examiner believes the paper is harder than last year's. The data shows candidates performed badly on question 3.
- ▶ Forecast grades are higher than they were last year.
- ▶ There has been a lot of change in the entry.
- ▶ All candidates also do coursework, and the coursework marks are a bit lower than last year.

What would you do?

Maintaining the awarding standard

- ▶ Setting and marking
- ▶ Grading - a wide body of evidence
 - ▶ Distributions of grades- compare centres who moved from June to March
 - ▶ Comparing candidates taking pairs of syllabuses between series
 - ▶ Consistency of thresholds over time
 - ▶ Tasks with unchanging thresholds – compare performance between these and other components
 - ▶ Examiner's professional judgement on challenge of assessment
 - ▶ Use of forecast grades

Marks and Results

- ▶ Example of marks conversion
- ▶ Enhanced results analysis

- ▶ Useful web link

<http://www.cambridgeinternational.org/images/206374-guide-to-converting-component-marks-into-syllabus-grades.pdf>

Cambridge Primary and Lower Secondary

- ▶ Checkpoint predictive validity
- ▶ The within-subject correlations were high for each subject (almost all above 0.70).
- ▶ In all subjects candidates achieving the very top Checkpoint score were most likely to get an A* at IGCSE.
- ▶ Higher Secondary Checkpoint scores were associated with an increased likelihood of achieving at least a grade C and at least a grade A at IGCSE.
- ▶ Science and Maths were better predictors than English.

Cambridge Primary and Lower Secondary

- ▶ Under review – consultation has been carried out
 - ▶ Security of Checkpoint tests – consideration of the ‘window’
 - ▶ Progression test availability
- ▶ Diagnostic testing service - not high stakes
- ▶ Model of production - item banked
- ▶ Model of deciding results is statistically based on item level data

Specific areas for discussion

- ▶ Core and extended qualifications
 - ▶ overlap of content
 - ▶ rationale
- ▶ Overseas examiners – legal considerations
- ▶ Models of assessing languages – First, Second, Foreign

Thank you
Any questions?

