



UNIVERSITY *of* CAMBRIDGE  
International Examinations

# Grading and Grade Review

Christine Cayley

University of Cambridge International Examinations



# Definitions

- **Grading** Process of deciding how many marks a candidate needs in order to obtain a certain grade
- **Grade Review** Process (before the issue of results) of re-marking the scripts of candidates thought most at risk of obtaining a wrong/inappropriate/ indefensible grade.



We also produce grades at  
component level, not just **syllabus**  
level.

So how many thresholds does that amount to?



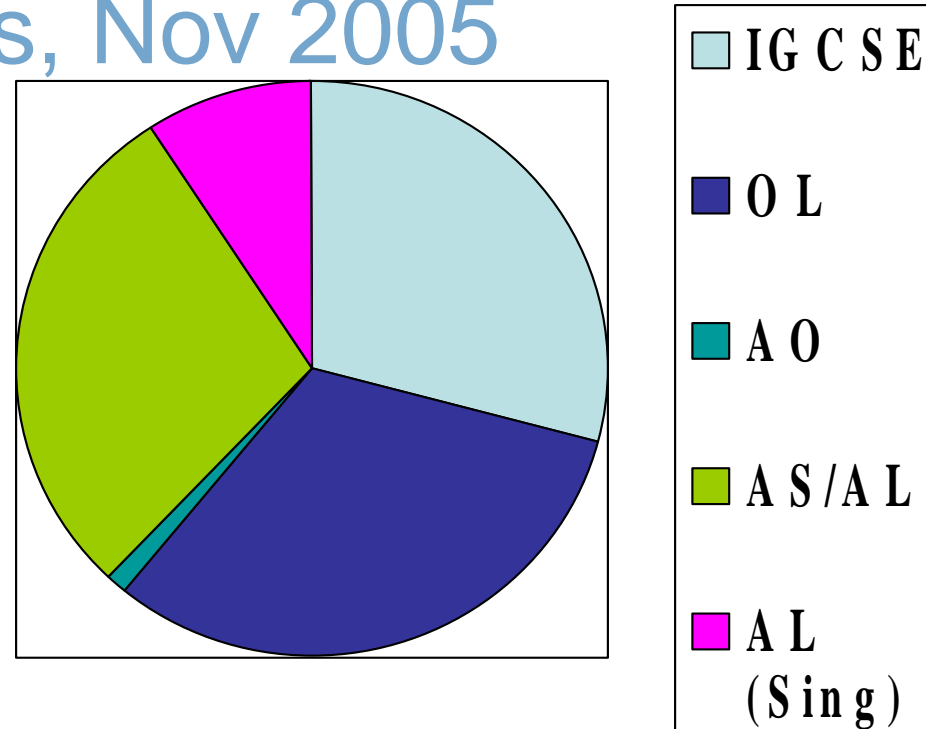
## Thresholds determined, November 2005 (1)

- IGCSE: **63 syllabuses** x say **4 components**  
(plus syllabus) x **8 thresholds** = **2,520**
- AS and A Level: **83 syllabuses** x say **4 components**  
(plus syllabus) x **6 thresholds** = **2,490**

## Number of thresholds, Nov 2005

- ~~between 218 PMs~~ (between 31 PMs)

IGCSE	2,520
OL	2,760
AO	104
AS/AL	2,490
AL (Sing)	800
Total	8,674
(between 31 PMs)	



## So how do we convert marks into grades?

First, some things we **DON'T** do:

- Set a pass mark that is the same for everything every year,  
e.g. 50% for a pass, 65% for a B, 80% for an A.
- This is what some people think we do. Would make life easy, but...



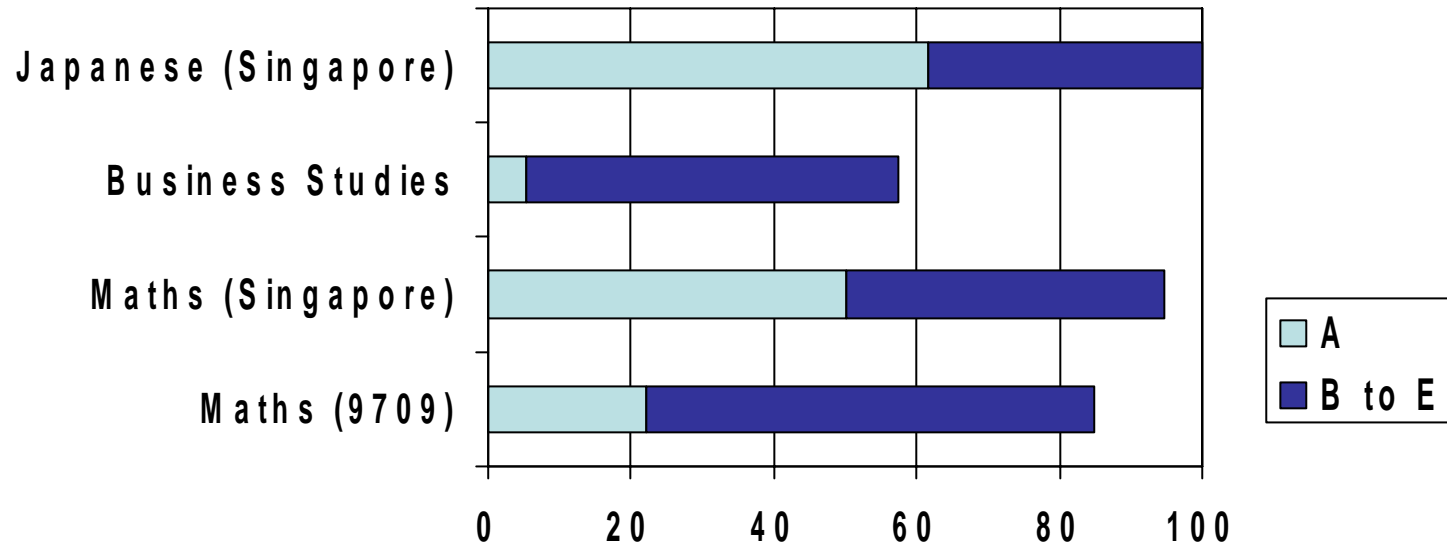
- Inappropriate across different subjects
- Inappropriate across different sessions/papers
- so we don't do that...

# How do we convert marks into grades?

- Another thing we could do, but **don't**, is to set a QUOTA of candidates who are going to pass at each grade, e.g. the top 10% will get an A, and so on.
- In 1970s guidelines for A Level said
  - A for top 10%
  - B for next 15%
  - C for next 10%
  - D for next 15%
  - E for next 20%.
- but...

- Inappropriate for different subjects
- Could be unfair for candidates who happen to be in a 'good' year
- Especially difficult for global examinations
- Countries come and go from syllabuses

## Cumulative % of candidates achieving Grade A or E in certain A Level syllabuses, November 2005





- If we're not maintaining the marks,
- and we're not maintaining the pass rate,
- what ARE we trying to maintain?



# THE STANDARD!





## The big question

- ***What is the lowest mark for which the performance this year can be deemed equivalent to that achieved by candidates at the lowest mark for the same grade last year?***



## The PEs' recommendations

- Each Principal Examiner makes recommendations for the answer to this question on Form SRS.
- They can state a range of marks /a tolerance.
- They have archive scripts.
- We remind them of last year's thresholds.



- HARDER paper should mean LOWER thresholds.
- BETTER candidature should mean HIGHER pass RATE.

## Judgemental thresholds (1)

- However, not for all the thresholds, just for certain ‘key’ (or ‘judgemental’) ones, i.e:
- A Level: A, B, E
- O (and AO) Level: A, C, E
- IGCSE A, C, E, F
- We sort these out and let arithmetic (and EPS) suggest the rest.

## Is that it?

- No, because we don't just do what PEs recommend.
- Why not?
  - Because sometimes their recommendations are inconsistent
  - Because they don't know what the outcome of their recommendations is going to be in terms of % of candidates getting through

## Basis of CIE grading

CIE grading is based on

- **statistical evidence**

combined with the

- **PEs' professional judgement**

as sorted out by

- **the professional judgement of the Product Manager.**

To do this, sometimes we have a Meeting.

# Setting Grade C threshold

Candidates had to do Papers 3, 4 and 6 if entered for the 'Extended' option

	max mark	C threshold
Paper 3	40	25
Paper 4	80	35
Paper 6	40	20
<b>Total</b>	160	<b>80</b>

## Setting Grade C threshold (2)

- So overall a candidate needed 80 out of 160 to be awarded a C?
- NO, not quite...

## Setting Grade C threshold (3)

- That would have been true if we had just used the ‘Indicator 1’ method of adding up the thresholds for the components, which means...
- .... adding up the thresholds for the components.
- But that’s too easy.

# Setting Grade C threshold

36.4% of candidates would get a C overall if we took a threshold of 80.  
% of candidates with C or above in different components (1)

0455/03	35.7
0455/04	42.2
0455/06	40.7
average	39.5
weighted average	40.2
MARK obtained by top 40.2%	77

# Putting components together

- So we took 77, not 80.
- 77 was the 'Indicator 2' mark obtained by the 'weighted average' method.
- We calculate both 'Indicator 1' and 'Indicator 2' and use the lower.



## Order of setting thresholds

- We set each key grade on each component for a syllabus,
- then set the key grades on the syllabus (or syllabus option),
- then check what we have done looking at various countries as well as the overall figures,
- then set the arithmetic boundaries,
- then check that they look right,



# Grading Confirmation

- then pass it over to the Standards and Projects Manager for a further check,
- and to CIE's Accountable Officer for another look.

# The complications

No use carrying forward last year's standard if we think:

- Subjects at the same level really were harder than one another

Or

- We are out of line with our UK equivalent

Or

- June is easier/harder than November

## Syllabus pairs (1)

- These are a way of comparing different subjects.
- We don't necessarily expect the same percentage pass rates in Economics and Principles of Accounting because the field of candidates will not be the same.
- However, we could isolate those candidates who did BOTH Economics and Principles of Accounting and, if we had enough candidates, we might expect them to do more or less equally well on the two.



## Syllabus pairs (2)

- If we found the same candidates did a lot better in Economics, we might take another pair, say Economics and Maths, and another, and so on, to check whether Economics was too easy.

# Grade Review

- This is a further check, after Grading, that candidates are likely to be given the right grade.
- It involves the re-marking of some candidates' scripts.

## Grade Review criteria

- They are likely to be re-marked if they are just below a grade threshold and
  - were forecast a significantly higher grade OR
  - did much better on one component than another OR
  - were marked by a ‘lingering doubt Examiner’ OR
  - are in a Centre that is doing much less well than forecast or than last year.
- We also sample all Examiners.

- So Grade Review is not like Grading.
- Senior Examiners come and mark scripts, and the Product Manager:
  - uses EPS to select the scripts in the first place,
  - cuts the coat according to the cloth,
  - gets the re-marking Examiners to follow up anything odd and be dynamic,
  - tries not to get grumpy.

# Why does it matter?

- It's the culmination of the work of CIE to which we all contribute.
  - Syllabus must be 'right'
  - Question Paper must be 'right'
  - Mark scheme must be 'right'
  - Marking must be 'right'
  - but Grading must be right too.



UNIVERSITY *of* CAMBRIDGE  
International Examinations

[www.cie.org.uk](http://www.cie.org.uk)