De-mystifying the role of the uniform mark in assessment practice: concepts, confusions and challenges

Elizabeth Gray OCR **and Stuart Shaw** CIE

Introduction

The search for an adequate conceptualisation of the Uniform Mark Scale (UMS) is a challenging one and it is clear that there is a need to broaden current discussions of the issues involved. This article marks an attempt to demystify the UMS; its conception and operation. Although the article assumes a basic appreciation of the terminology and processes associated with the examination system, it explicates through a number of case study scenarios, the contexts in which it is appropriate to employ UMS, describes any necessary computations arising from different specifications and assessment scenarios, and addresses some of the potential challenges posed by the calculation of grades for unitised specifications. A specification here refers to a comprehensive description of a qualification and includes both obligatory and optional features: content, and any performance requirements. If a specification is unitised, the constituent units can be separately delivered, assessed and certificated.

Having a clear and well articulated position on the underlying theory of UMS is necessary to demonstrate transparency with regard to the estimation of aggregate performance on unitised assessments and to support any claims we wish to make about the reporting process.

It is hoped that the issues addressed here will make a positive contribution to the widening nature of the UMS debate (both within and beyond Cambridge Assessment) more generally, and of the understanding, operation and employment of UMS, in particular.

Underlying rationale for a Uniform Mark Scale

Educational assessments are currently delivered as either non-unitised specifications or as unitised ones. By non-unitised, we mean that candidates take the various components (which may be written papers, coursework or controlled assessment) that make up the specification in the same session or administration. Following any examiner or moderator scaling adjustments, the marks for candidates are aggregated to give a total mark for the entire specification: this defines the specification (also referred to as syllabus by Cambridge International Examinations, CIE) grade. The purpose of the grading process is to determine, the lowest mark for which the performance in the current administration can be deemed equivalent to that achieved by candidates at the lowest mark for the same grade for the last administration. Grading is undertaken for each key threshold on each component and each specification option. Generally the grading process attempts to involve comparisons with the standards set in previous sessions, but occasionally the process is one of standard setting for a new specification. Using the grade boundaries established by an awarding committee, a candidate's specification grades are subsequently determined from the total marks.

Unitised assessments, however, allow the candidate to take the unit assessments (the smallest part of the specification formally reported and certificated) on different occasions. Unitised assessments may exhibit variation in their respective levels of difficulty over time. Where this happens the grade boundaries for a January unit assessment, say, may be slightly different from those set for the corresponding May/June unit assessment. It is crucial, therefore, that a mechanism be implemented for mapping different marks awarded on different occasions onto some common scale such that the differing marks constitute the same value when aggregated to give an overall grade.

Issues relating to Aggregation

Aggregation is 'the process of combining (by summation or other agreed procedure) the marks or other units of credit awarded through an assessment scheme' (QCA Code of Practice, 2007, p. 65). Aggregation issues are a source of constant debate within the public examination area and Thomson (1992) provides a good description of the issues relating to methods which seek to combine raw marks (the marks originally awarded when assessed) of units achieved at different times and with different grade boundaries.

Potentially, there are a number of methods for combining raw marks of units achieved at different times. According to Thomson, many give rise to one of two types of anomaly. In Type I anomalies, two candidates with the same grade profile across four units receive different subject grades. For example, 'abbd' = B; and, 'abbd' = C. A special case of this is a candidate who obtains the same grade for all units, but obtains a different subject grade, for example, 'bbbb' = A. In Type II anomalies, two candidates with a different profile obtain the same grade. For example, 'abbc' = B; and, 'aabb' = B[1].

Different methods of aggregation give rise to different instances of these anomalies. Unless there is a very crude system of assigning a point to a grade, all methods will result in at least some Type II anomalies, and many in Type I. One of the reasons for the choice of uniform marks for aggregating unitised schemes is, therefore, that the instances of anomalies can be reduced if the conversion is suitably chosen (Thomson, ibid).

With the introduction of unitised schemes of assessment, GCE became wholly unitised in 2001/2002 although there were modular forms of general qualification assessments before then, and these add an additional complexity to the aggregation process because units may be taken within the duration of a course of study, not just terminally. In order to be fair to these candidates when a specification grade is calculated raw marks cannot be used. The reason is perhaps best illustrated by use of an example:

1 Clearly different grade profiles can lead to the same or different overall outcomes, some of which may be counter-intuitive.

Imagine a candidate takes a unit twice and achieves 72 raw marks in the first instance and 68 on re-sitting. On the first occasion the A boundary is set at 73 whilst on the second it is set at 67. In the case of the re-sit, the candidate gains a higher grade with a lower mark than in the first examination and the 'value' of raw marks is not the same for the two examinations.

An elementary approach for resolving this difficulty might be to award grades only to candidates on each unit. Unit grades would then be assigned numerical points (A*=9, A=8 ... U=1) and then a simple addition of points would provide a total for the specification. There are two distinct disadvantages with such a rudimentary method:

1. this approach would not discriminate between weak, adequate and strong performances within the same grade, in other words marks provide more detailed information than grades; also
2. problems would arise where units were unequally weighted. The weighting of an assessment is its overall contribution to the total or aggregate assessment. For example, if a unit is weighted at 35%, the unit accounts for 35% of the total assessment. In this case, a scaling factor would need to be applied to the raw marks in order to give them the appropriate weighting.

In order to obviate these shortcomings, a segmented linear scaling methodology is used. Such deficiencies are thus re-mediated by adoption of a procedure which utilises a common mark or standardised scale. A standardised mark is the result of a transformation of raw marks which provides a measure of relative standing in a group and allows comparisons of raw marks from different distributions (Davies *et al*. 1999, p. 186). A common scale[2] has the advantage of affording greater credit to candidates who have achieved higher marks within a grade and legislating for unequal unit weightings by setting a uniform mark scale for the unit which reflects its weighting in the specification.

In outline, raw marks are mapped on to a scale which takes into account the value of the raw mark. This scale is known as the uniform (or standardised) mark scale. Here, if the unit is worth 100 UMS then the A boundary is 80 UMS (using the usual GCE UMS). Taking the example introduced earlier: on occasion 1, the boundary of 73 raw marks would map to the UMS boundary of 80 and the candidate would get 79 uniform marks. On occasion 2, the boundary 67 raw marks would map to the UMS boundary of 80 uniform marks and the candidate would get 81 uniform marks. In this way the value of the raw mark, in terms of the grade it would earn, and the quality of that grade, are preserved.

The important point to note here is that uniform mark scales remain the same throughout the lifetime of the specification and, particularly, from one session to another. This means that the grade which a candidate receives and the position of the raw mark within the grade bandwidth (i.e. the marks between the two grade boundaries within which the raw mark sits) will always convert to the same uniform mark irrespective of the actual raw mark and the raw mark boundaries.

Thus a *uniform mark* is used when units of an assessment can be taken on different occasions during a course of study and is a mark on a standard scale which indicates a candidate's performance. A *uniform mark scale* is a means of achieving parity between alternative units in specifications and functions to effectively smooth out the small variations in the demand[3] of the assessment units sat by candidates during their GCE and GCSE studies. *Uniform mark boundaries* for unit and specification conversions remain the same for the lifetime of the specification.

It is a requirement of the QCA that aggregate marks from a unitised GCE or GCSE or staged tests should be computed on the basis of a UMS: 'Uniform marks for each unit must be calculated in such a way as to maintain the candidates' relative position between the raw grade boundaries. Each unit must be reported in uniform marks. Uniform marks for individual assessment units are added to generate a final grade for the qualification as a whole' (QCA Code of Practice, 2007, p. 56). The requirement for converting raw marks to uniform marks for the purposes of aggregation facilitates fairness of the specification outcomes.

The relationships between uniform marks and grades are shown in the relevant GCE and GCSE specifications and uniform marks and unit grade results are distributed to centres in the Cumulative Specification Results Report and to candidates in their Statement of Results.

We now turn our attention to how a uniform mark is calculated and in the computation process begin to appreciate some of the potential challenges unearthed by aggregation, highlighting some of the relative merits and de-merits of unitised schemes of assessment.

Conversion of raw to uniform marks

We have seen that a candidate's raw marks are mapped onto a scale which is invariant for the lifetime of the specification. The conversion of raw to uniform marks is dependent on the grade boundaries on the occasion when the raw mark was achieved.

The uniform mark scale will have been determined when the specification was originally accredited. So, for each unit, uniform mark maxima and grade boundaries are pre-set. In order to preserve the value of a raw mark, there should be a one-to-one, linear mapping of the uniform marks with the raw marks. In this instance, the boundaries are reasonably spaced and there are no issues relating to effects which manifest at the extremes of the mark distribution. This is illustrated in Figure 1 which assumes there are 100 raw marks. In this case, the GCSE A* boundary of 90 uniform marks will coincide with the raw mark boundary of 90 raw marks.

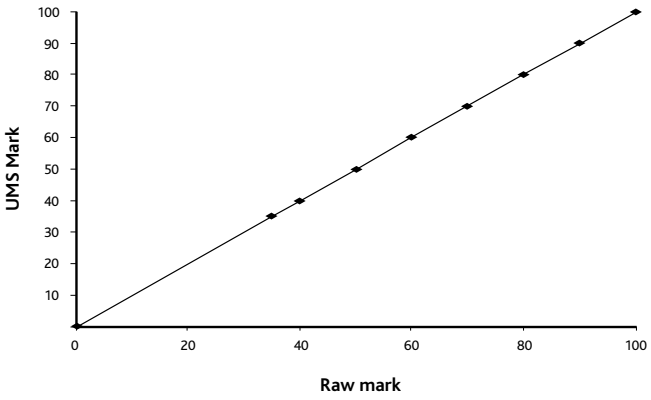

Figure 1: Ideal raw mark/uniform mark relationship

2 The simplest form of common scale would award one point per grade for equally weighted units which would not differentiate between candidates within a grade. It is rarely used in general qualifications for this reason.

3 Demand in this context is defined by awarders when setting grade boundaries. This is a judgement made in the presence of performance (candidates' work) and statistical evidence.

Figure 2 depicts another simple raw to uniform mark conversion which demonstrates linearity between grades A and E. However, above and below the two end points the conversion factor changes as is shown by the change in the slope of the line. This is because grades A and E are recommended at the grade award and intermediate grades are interpolated maintaining (to within a mark) the linear relationship of raw marks to uniform marks. However, unless A and E are chosen so that the mapping would continue the straight line between A and the maximum and E to zero then the line will consist of three segments. In fact, in Figure 2, raw marks above A and below E would be worth a smaller number of uniform marks than each individual raw mark between the A and E boundaries. The conversion factor from raw to uniform marks is smaller above the A boundary and below the E boundary than between the A and E boundaries.

Figure 2: A simple UMS conversion

In order to convert from a raw to a uniform mark:

1. Determine what grade the raw mark would indicate based on the raw mark boundaries for the unit.
2. Calculate the number of raw marks in the raw mark grade bandwidth.
3. Calculate the number of uniform marks in the uniform mark grade bandwidth. This number will be based on the weighting of the unit and the maximum UMS for the specification. All uniform marks between the top and bottom of the grade range will be the same, but may differ above the top and below the bottom grade.
4. Calculate a conversion factor (the number found in Step 3 divided by the number found in Step 2).
5. Find the difference between the lower raw mark grade boundary and the raw mark for conversion and multiply it by the conversion factor (from step 4).
6. Add the number found in Step 5 to the lower uniform mark grade boundary.

We can see how this process is operationalised in a live context by giving consideration to the following scenario:

Imagine a GCE candidate gains a raw mark of 49, the raw mark grade boundary for grade C is 47 and, the raw mark boundary for B is 53. Additionally, the UMS B boundary is 70 and the UMS C boundary is 60.

1. *The candidate has 49 raw marks which lie between the C and B boundaries so the candidate has a grade of C.*
2. *The raw mark B boundary is 53 and the raw mark C boundary is 47 giving 6 marks in the grade bandwidth.*
3. *The UMS B boundary is 70 and the C boundary 60, i.e. there are 10 uniform marks in the uniform mark grade band width (range).*
4. *The conversion factor is found by dividing the number of uniform marks in the range (10) by the number of marks in the grade bandwidth (6) i.e. 10/6= 1.7.*
5. *The candidate's raw mark is 49 and the lower raw mark grade boundary is 47. So the number of marks above the grade boundary is 49–47 = 2. The result of this calculation is multiplied by the conversion factor, i.e. 2 * 1.7 = 3 (rounded).*
6. *The result of step 5, i.e. 3, is added to the lower uniform mark grade boundary, i.e. 60. Therefore, the candidate's total uniform mark is 60 + 3 = 63.*

There are many different UMS which can be constructed ranging from a simple point per grade to the regime currently used for GCE and unitised GCSE specifications. It has been shown that matching uniform marks to raw marks as far as possible reduces the number of anomalies. Too short a scale and approximation and loss of information gives rise to a reduction in percentage at the top end of the grade range and considerable unfairness because the quality of the raw mark is not taken into account. Too long a scale implies spurious discrimination and has proved difficult to explain to centres.

The effect of approximating the raw mark too much is demonstrated in the next example:

In a three unit GCSE, a candidate achieves 74, 89 and 68 marks respectively. In this very simple example, the candidate would get an A grade from the raw mark or from the UMS (using 100 UMS for each unit with boundaries at 90%, 80% and so on) but not using a points conversion. This is because the quality of the grades – one just below the A boundary and the other units just below A – is not recognised.*

Table 1: Raw marks to points conversion

	Unit 1	*Unit 2*	*Unit 3*	*Total*
Raw max	100	100	100	300
Raw A*	80	90	69	239
Raw A	75	81	64	220
Raw B	70	72	59	201
Point max	8	8	8	24
Points A*	8	8	8	24
Points A	7	7	7	21
Points B	6	6	6	18
Candidate raw	74	89	68	231
Candidate points	6	7	7	20
Candidate UMS	78	89	88	255

Uniform grade boundaries in GCSE and GCE specifications

Uniform grade boundaries in GCSE and GCE specifications are established by inter-awarding body agreement. Table 2 shows the mark grade boundaries as percentages of the maximum uniform mark for the unit (or module).

Most GCE subjects are currently based on a 600 uniform mark total. Therefore, the uniform mark grade boundaries for an Advanced

Table 2 Uniform grade boundaries: GCSE and GCE

	GCSE							
Grade	*A**	*A*	*B*	*C*	*D*	*E*	*F*	*G*
%	90	80	70	60	50	40	30	20

	GCE				
Grade	*A*	*B*	*C*	*D*	*E*
%	80	70	60	50	40

specification are A: 480 (= 80% of 600), B: 420 (= 70% of 600), C: 360, D: 300, E: 240. For an evenly balanced scheme of six, equally weighted units, each unit attracts a maximum mark of 100 uniform marks after conversion, with 80 for an A, 70 for B and so on. This gives an A range of 20% of the uniform mark range, with the other pass grades all having the raw grade range mapped on to 10 marks. If the units are not equally weighted or totalling six in number, or both, the UMS for each unit is usually calculated to be in the proportion of that unit of 600, with the boundaries set accordingly, so that in all such cases there will still be greater compensation for an A than any other grade. Table 3 shows this more explicitly for the commonest weightings of 15%, 16.7% and 20%.

Table 3: UMS for GCE specifications

Grade	*Percentage of maximum UMS*	*Specification*	*15% weighted unit*	*16.7% weighted unit*	*20% weighted unit*
Max	100	600	90	100	120
A	80	480	72	80	96
B	70	420	63	70	84
C	60	360	54	60	72
D	50	300	45	50	60
E	40	240	36	40	48

Whatever combination of weighted units are added together (provided the total weighting is 100%), the percentage of marks at each grade boundary will be the same. Therefore, five 20% weighted units will equate, in percentage terms to six 16.7 % weighted units or any other combination. The specification boundary marks will always be the same. In fact, there are almost always six units in a GCE examination, but the weightings are in a variety of combinations.

In September 2008 new GCE courses will start with the first candidates taking A2 examinations in 2010. Most of these will consist of 4 units with a total uniform mark out of 400, although percentages will remain unchanged, that is, 80% of the uniform marks available will determine the A boundary and 40% the E. However, a major challenge will be the introduction of the new A* grade. This is to be awarded to candidates gaining an A grade overall and 90% of the uniform marks available on the A2 units (the second half of the A level). Ensuring fairness and comparability for A* candidates will depend critically on the conversions above A.

Conversions are similar for untiered GCSE assessments, with 90% of the available range assigned to A* with 10% grade bandwidths down to 20% for a G. Maximum uniform marks are not prescribed and are usually chosen as a best fit with the assessment structure. Although the same UMS applies for tiered specifications there are some differences because of the tiers. The maximum uniform mark for a foundation tier unit will be one uniform mark below the B boundary, and the allowed E on the higher tier is set at half the uniform mark grade bandwidth below D (Table 4).

Table 4: Uniform mark boundaries for a 100 ums unit

	Max	*A**	*A*	*B*	*C*	*D*	*E*	*F*	*G*
Untiered	100	90	80	70	60	50	40	30	20
H tier	100	90	80	70	60	50	45		
F tier	69				60	50	40	30	20

Because of the more complex grading regime for GCSE tiered specifications and the additional judgemental boundaries, uniform mark conversions can be more complex, not least because they are potentially different for each unitised GCSE.

It is also important to note that grade boundaries on a uniform mark scale are all the same percentage of the maximum mark. Thus, for GCE assessments, 80% of the maximum mark at both unit and specification level gives the A grade boundary and 40% the E. If it were not so it would be impossible to combine units with different weightings and still maintain the same specification grade boundaries. Table 5 exemplifies the issue.

Table 5: Points conversions for differentially weighted units

Grade	*25%*	*50%*	*75%*	*2@25% +1@50%*	*1@25% +1@75%*
A	6	11	16	23	22
B	5	9	13	19	18
C	4	7	10	15	14
D	3	5	7	11	10
E	2	3	4	7	6
U	1	1	1	3	2

In this very trivial example, and with the points as indicated, the aggregation of differently weighted units leads to different maximum and grade boundary marks. This would be possible to control within a specification which dictated the weighting of each unit, though somewhat confusing; but in a specification, various combinations of units are permitted and such a points regime would be unacceptable.

In all unitised general specifications a grade E is half the value of a grade A at 40% and 80% of the maximum UMS respectively. This relationship is important because a change would affect the basic characteristics of the conversions.

Uniform Mark Scales: challenges and confusions

Uniform marks are not without their difficulties although a range of differing stratagems have been used to overcome the worst. One of the basic issues related to uniform mark use is the maintenance of the value of each raw mark within a unit. No distinction is made, on an assessment's raw mark scale, as to the value of each raw mark and they are, for the purposes of aggregation, all deemed to be of the same value.

The same may not be true after conversion to uniform marks because of the nature of the conversion. If the conversion line is not strictly linear, even if the discontinuities only affect the extreme grades, the consequences are not only undesirable, but also difficult to explain.

In Figure 2 above we see that the conversion line is discontinuous and the conversion rate differs depending on the position of the raw mark relative to the grade boundaries. Grading rules will almost always lead to a line which is segmented, that is, piecewise continuous. For GCE there are two judgemental grades, A and E, and if they are not set exactly at 40% and 80% of the raw mark scale then the line will be discontinuous. Intermediate grades are arithmetic (i.e. determined mathematically), and if the number of marks between A and E is not exactly divisible by 4 (the number of intervening bandwidths) there will also be discontinuities in the line between A and E. For GCSE there are more judgemental boundaries, A, C and F (with D on the higher tier of a tiered unit) and more arithmetic boundaries to be set with a greater likelihood of several conversion factors for the raw marks being applied.

In order to minimise these differences Cambridge Assessment has long had a policy of targeting grade boundaries to align with the UMS scale. So, in writing GCE papers, for example, the question setter will aim for a minimum 'A' performance at around 80% of the raw marks and a minimum 'E' performance at about 40% of the raw marks. The problem is even greater with GCSE assessment because there are more boundaries to be set and potentially different conversion rates between grades which can lead to unpredictable consequences. For tiered specifications with a continuous scale through the tiers it is impossible to set targeted boundaries which will lead to a continuous line, the aim there is to minimise the discontinuities as far as possible.

One of the problems that arose as a result of unequal conversion rates in earlier modular schemes was a reduction in the expected number of high grades. Part of this was due to the lack of consideration of the effect of UMS conversions when mark schemes were devised and often much of the compensating power of a 20 mark UMS A range was lost because raw mark ranges above A were too long, or were not fully utilised. The other reason was the effect of regression. Awarding bodies have addressed the former by the use of a capping mechanism.

Figure 3 shows the effect of capping. Continuing the line showing the conversion from raw to uniform marks between A and E (the dash-dotted line) so that the raw marks above A retain their value, it can be seen that the line reaches the maximum uniform mark before it reaches the maximum raw mark. This is the effect of capping. Without this intervention, candidates gaining raw marks above the A boundary would have conversions based on the dotted line which would not give as much value to their raw marks.

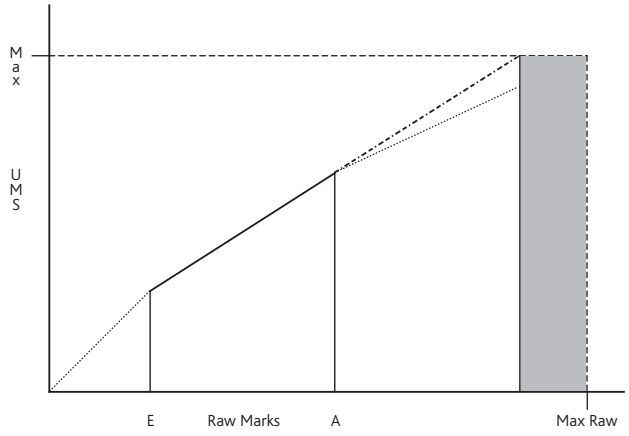

Figure 3: The effect of capping

Candidates in the shaded part of Figure 3, whilst not attaining maximum raw marks, will receive maximum uniform marks. The issues relating to capping can be best illustrated through the following example:

A two tiered GCSE unit has raw marks of 60 in each tier. The maximum uniform mark for the unit is 100. The A, C and D boundaries on the higher tier are 42, 31 and 26 respectively; the foundation tier C and F boundaries are 47 and 28 respectively. The maximum uniform mark for the foundation tier is 69 (one below B). The allowed E will be 45 i.e. 50 – (60–50))/2.

- *On the Higher tier, the B boundary is 36, which makes the A*boundary 42. The tier will be capped at 48 raw marks. All candidates scoring 48 and above on this tier will receive 100 uniform marks. The allowed E is at 23 raw marks and this will be mapped to 45 uniform marks.*
- *On the Foundation tier, the D boundary is 40, so this tier will be capped at 54 and all candidates gaining 54 or more raw marks will receive 69 uniform marks.*

Capping can, however, have undesirable consequences if the full raw mark range has been used. Usually the reason for a low A boundary is that high marks are unattainable so the fact that candidates with marks not on the maximum, but close to the maximum, will be given full uniform marks is not an issue. This might very well happen with the introduction of 'stretch and challenge' questions ('stretch and challenge' questions constitute a potentially promising solution to the issue of high achievement recognition although they are not without their challenges). However, if there is a low A boundary, but the full mark range has been utilised, there may well be significant numbers of candidates on the maximum uniform mark who have achieved very much less than the maximum raw mark.

Capping will also occur when there is a high E boundary even if A is about the 80% mark because twice the A/B distance will be shorter than the maximum raw mark. There is also another issue with a high E boundary. Conversion rates below E will be less than one and the effects of greater value being given to raw marks above E can lead to an unexpected increase in numbers passing the unit. For this reason, for GCE units, a notional N grade has been introduced to ensure that conversion through the E boundary is linear.

A second factor affecting marks at the top and bottom of the grade range is regression which is more accurately known as *attenuation of variance*. This is due to a bunching of marks on aggregation resulting in a reduction in the percentage of candidates gaining the top grades compared with the mean percentage taken over all units. The reverse effect is seen at the bottom of the grade range with an increase in the percentage of E grades. Neither effect is as a result of UMS conversions, although it might be exacerbated as described above.

One of the criticisms which attaches to the UMS method of aggregation is its *invariance*. Specification (and unit) boundaries are pre-defined and thus not open to 'statistical and technical' adjustment post hoc such as may be found with linear schemes. If such variation year-on-year were allowed, then this could give rise to a third type of error. Candidates with the same uniform mark total could be getting different specification grades from year to year. Since raw grade boundaries are set

to allow for differences in demand, the point about the UMS conversion is that this differential has been allowed for. Looked at from a raw mark perspective, if specification uniform mark grade boundaries are allowed to fluctuate (but not unit conversions), then the relationship of raw unit boundaries to that final total will vary. Even calculating a regression allowance of UMS marks would lead to year-on-year anomalies because candidates on what were ostensibly equivalent marks could achieve different grades purely because of the company they keep even though much of their assessment might be common.

Conclusions

This article has attempted to explain the underlying rationale for the employment of uniform marks: their conception, their computations; and their effect on a range of aggregations. The principal motivation for using the uniform mark scale relates to the structure of regulation for GCE specifications and of choice for GCSE development unit based.

The relative strengths and shortcomings of using uniform marks for unitised schemes of assessment are both multiform and various. Unitised schemes are flexible, enhance overall performance (although some would say unfairly because of the provision for re-sits) and enable weaker candidates to show what they know, understand and can do because the learning approach is both incremental and developmental: learners have greater control regarding choice of assessment without undue reliance on terminal assessment. Unitised assessments are manageable, formative and can be delivered at the point of learning within the programme of study. Additionally, GCE and GCSE are similar in basic structure with units employing credit ratings which have the potential to be used in a National Qualifications Framework and as part of the Additional and Specialised Learning in the Diploma.

Conversely, there is a prevailing belief that unitisation can lead to increased testing and, therefore, to a concomitant increase in the burden of assessment. More disturbingly, there exists a public perception that unitised schemes are easier, largely due to the re-sit policy. From a cognitive maturation perspective, it is also held that some candidates who take unitised assessments may forget that part of the curriculum very readily. This has led to synoptic assessment in GCE specifications and terminal rules for the new GCSE developments.[4] In terms of their interpretation, evidence would suggest that centres find it difficult to read and comprehend UMS data. We have seen that there are problems when there are discontinuities in the conversion rates which have led to the generation of some additional rules to maintain conversion parity.

Whatever the arguments, the UMS system has stood the test of time (it was first introduced as a mechanism for aggregating GCE specifications in the late 1980s) and, with the modifications described, seems to work well. There are concerns that with the new A levels and the introduction of 'stretch and challenge' questions it will be difficult to target grades as precisely as is achieved with the current GCEs with the inevitable consequences of low grade A and, possibly, E boundaries. GCE A* is another complication because its achievement is crucially dependent on the amount of capping there is in the specification. But until another, more effective, system is devised for aggregation, uniform marks are likely to remain.

References

Davies, A., Brown, A., Elder, C., Hill, K., Lumley, T. and McNamara, T. (1999). *Dictionary of Language Testing, Studies in Language Testing 7*. Cambridge: UCLES and Cambridge University Press.

Greatorex, J. and Malacova, E. (2006). Can different teaching strategies or methods of preparing pupils lead to greater improvements from GCSE to A level performance? *Research Papers in Education*, **21**, 3, 255–294.

Qualifications and Curriculum Authority (2000). *Arrangements for the statutory regulations of external qualifications in England, Wales and Northern Ireland*. London: QCA.

Qualifications and Curriculum Authority (2007). *GCSE, GCE, GNVQ and AEA Code of Practice*. London: QCA.

Patrick, H. (2003). Synoptic Assessment: A report for QCA. Available at http://www.ofqual.gov.uk/files/synoptic_assessment-_report_for_qca_pdf_05_1620.pdf

Pollitt, A., Ahmed, A., and Crisp, V. (2007). The demands of examination syllabuses and question papers. In: P. Newton, J-A Baird, H. Goldstein, H. Patrick, and P. Tymms (Eds.), *Techniques for Monitoring the Comparability of Examination Standards*. London: Qualifications and Curriculum Authority.

Stevens, J. (2002). The demands of synoptic assessment in the new English literature A level. *The Use of English*, **53**, 2, 97–107.

Thomson, D. G. (1992). *Grading Modular Curricula*. Cambridge: Midland Examining Group.

4 QCA has defined synoptic assessment as follows (QCA, 2000): A form of assessment which tests candidates' understanding of the connections between the different elements of a subject. See also Patrick, H (2003) and Greatorex and Malacova, (2006).

CAMBRIDGE INTERNATIONAL EXAMINATIONS

The CIE Research Agenda

Stuart Shaw CIE

Introduction

Cambridge Assessment has long devoted attention to assessment research. As part of its on-going commitment to examination quality, Cambridge International Examinations (CIE) has developed and established a unit dedicated to research. Although small, the team is responsible for a variety of research activities ranging from routine operational procedures in support of the quality of assessment processes to more full-scale experimental investigations whose purpose is to inform and improve on those operational procedures.

The research unit is responsible for three main areas of activity:

- **Routine operational analysis** concerning the management cycle of all CIE assessments, including the examination production, conduct, marking and awarding, and post-examination appraisal.