Cite as: Little, C. and Jones, K. (2008), Assessment of University-entrance level mathematics in England: an analysis of key influences on the evolution of the qualification during the period 1951-2001. Paper presented at the *11th International Congress on Mathematical Education* (ICME-11), Monterrey, Mexico, 6-13 July 2008.

Assessment of University-entrance level mathematics in England: an analysis of key influences on the evolution of the qualification during the period 1951-2001

Chris Little and Keith Jones, University of Southampton, UK

Abstract

This paper tracks the evolution of the mathematics qualification, usually called GCE A-level, originally developed and commonly used for University-entrance in England. The analysis focuses on key influences on the evolution of the qualification during the period 1951-2001, looking at how, and by whom, change was initiated, and what provided the catalysts for changes to syllabuses and schemes of assessment. The analysis describes the infrastructure of the particular qualification, identifies a stakeholder model, and proposes catalysts for change in terms of issues relating to syllabus content, standards, validity and reliability. The analysis suggests that a significant feature of the qualification framework in England is the 'market' created by having multiple examining agencies delivering a single qualification. Whilst this does create problems of comparability, and can lead to pressures which are not wholly educationally desirable, the market does create a dynamic for change which a public examination monopoly might not readily achieve.

1. Introduction

The influence of formal public examination systems on the classroom is readily acknowledged by all teachers, and has not always been considered to be benign. The extensive research literature has criticised the narrowing and stultifying effects that summative examinations can have on learning, something touched upon by many of the papers in the study on assessment in mathematics carried out by the *International Commission on Mathematical Instruction* (Niss 1993a, b). Yet public examinations are a fact of life in most educational systems, and high-stakes school leaving examinations have become crucial rite of passage to be negotiated by students, giving them the license to progress to the next stage of their lives, in employment, vocational training or higher education.

Griffiths and Howson (1974) trace the roots of public examinations back to the mandarinate of ancient China. Systems of public examinations have evolved in different countries in disparate ways. Many countries use a model based upon the Prussian *Abitur* or French *Baccalaureate*, both developed towards the end of the 18th century. In England[1], the General Certificate of Education at Advanced level (shortened to GCE A-level) is the longest standing qualification (the Ordinary or 'O' level was superseded in 1988). This GCE A-level qualification was developed in 1951 out of the Higher School Certificate, whose origins lie in university matriculation examinations, developed initially by the universities of Oxford, Cambridge and London. Public examination systems in England grew largely through pressure on higher education to provide a fairer qualification for selecting students for mathematics courses in universities.

In terms of the A-level Mathematics qualification in England, this has developed in response to societal, technological and cultural changes. The 'modern mathematics' movement of the 1960s, which spawned high profile projects such as the *School Mathematics Study Group* (SMSG) in the US and the *School Mathematics Project* (SMP) in the UK, developed from the expansion in the industrial applications of mathematics such as statistics, operational research, linear programming and numerical analysis, coupled with a somewhat contradictory movement towards the inclusion of more abstract mathematics such as functions, matrices, vectors, group theory and linear algebra.

In the 1970s and 1980s, the qualification was required to respond to the move in England away from a bipartite system of grammar and secondary modern schools to a unitary system of comprehensive schools, and subsequent changes in the national qualifications at age 16, including the introduction of the General Certificate of Secondary Education (GCSE). The endpoints of pre-16 courses define the

[1] In the UK, Scotland has a separate educational system to England, Wales and Northern Ireland, and is excluded for this discussion. For the sake of simplicity, we use the term 'England' to refer to 'England, Wales and Northern Ireland'.

starting points of post-16 courses, and the resulting differentiation in outcomes at 16 created the pressure for curriculum development and change in GCE A-level mathematics.

The 1990s saw the development of alternative forms of assessment, such as practical and investigational project work and comprehension papers. This has recently been reversed, and assessment of the GCE has returned almost universally to relying on externally set, timed written papers. The influence of calculation and computing technology has also changed the nature of assessment at this level.

The need to 'control' the outcomes of the educational system has led to the development of a national curriculum for England and greater, more centralised political control of the school curriculum. Throughout this period of development, there has been a continuing debate on 'standards', prompted by the conflicting demands of expanding the number of students accessing higher education, whilst maintaining the quality of matriculation qualifications.

The questions this paper addresses are how one might model the evolution of a qualification like A-level Mathematics; how and by whom change is initiated; and what has catalyses the changes to syllabuses and schemes of assessment. In modelling the evolution of the A-level Mathematics qualification in England, the paper focuses on key influences on the evolution of the qualification during the period 1951-2001, examines how, and by whom, change was initiated, and identifies what provided the catalysts for changes to syllabuses and schemes of assessment. Such an account is likely to prove useful when comparing with other systems of qualification across the world.

2. Definition of terms.

In common parlance, the meaning of terms such as 'examination', 'syllabus', 'assessment' and 'qualification' are used somewhat loosely. Wel use the following nomenclature when describing public examination systems.

A *qualification* is a nationally sanctioned standard of accomplishment in a subject or subjects which, when achieved, confers status and recognition of an agreed level of attainment of knowledge, skills and understanding in the subject(s).

Examples: International *Baccalaureate*, *Abitur*, GCE Advanced level.

In order to achieve a qualification in a curriculum subject, candidates are assessed on a specified subset of knowledge and skills in the subject. These are communicated in a *syllabus*, a printed document developed by an *examining body*. More than one syllabus may be developed leading to the same qualification, and examining bodies may themselves develop more than one syllabus in a subject.

Examples: in England[2], there are five examining bodies (or 'groups'), commonly referred to by their acronyms: OCR, AQA, Edexcel, WJEC and CCEA, currently offering 6 syllabuses for GCE A-level Mathematics.

The syllabus specifies (a) *the assessment aims and objectives*, (b) *the mathematical content* (c) *the scheme of assessment.*

(a) The *assessment aims and objectives* describe in general terms the skills and abilities which are to be tested in achieving the qualification.

Examples of *aims*: develop understanding of mathematics and mathematical processes, develop the ability to reason logically, use mathematics as a means of communication, acquire skills in the use of calculators and computers.

Examples of *objectives*: Recall, select and use knowledge of standard mathematical models to represent situations in the real world; construct rigorous mathematical arguments.

(b) The *mathematical content* details the mathematical concepts, results, calculations and processes which will be tested by the scheme of assessment.

Examples: The product rule for differentiation, the exponential function, solving trigonometric equations, the compound angle formula, etc.

[2] See footnote 1

(c) The *scheme of assessment* specifies the assessment methods used by the syllabus. During the period of its existence, A-level Mathematics syllabuses have used a variety of assessment tools:

- *Timed, written papers* are the most commonly used tool, and familiar to all. These usually comprise a number of questions to be solved, varying in length from a few minutes to about 20 minutes; however, other types of paper, such as comprehension papers, are used. The papers are taken at fixed dates, currently in January or June/July, and are externally set and marked.
- *Coursework* is the term used in the UK for other pieces of student work used as part of the scheme of assessment. These tasks are set by the teacher or examining body, and are completed as part of the course, either at home or in the classroom. Coursework can vary in length and open-endedness, and may be used to test aspects of the construct which are difficult to test using timed written papers, such as mathematical modelling, investigative skills, or use of computers.

Finally, examining bodies publish *assessment materials*, comprising past timed written papers with mark schemes (although in the past these were not released to schools), and supplementary guidance for teachers on coursework.

Whilst the terms outlined above may differ in usage in other countries, these will be the terms used in this paper to describe a public examination system. Diagrammatically, the structure may be illustrated as in Fig. 1

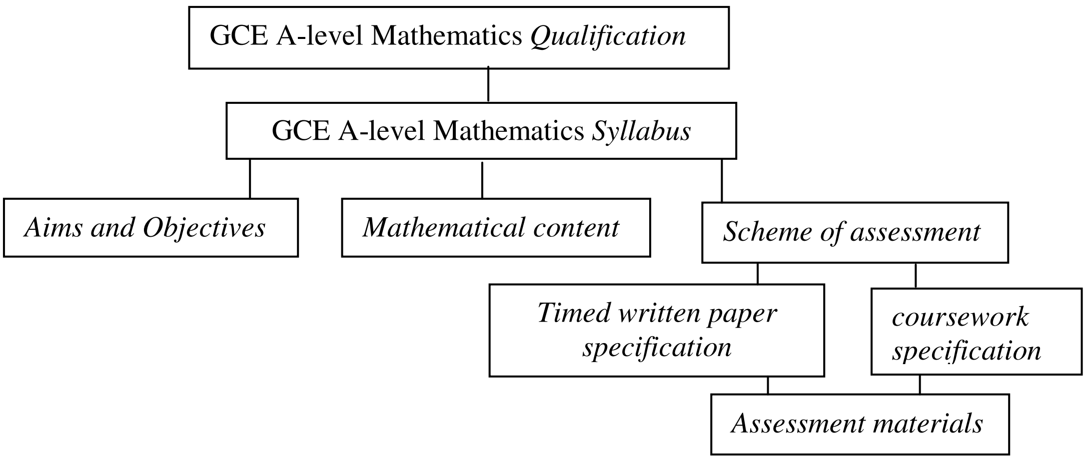

Fig. 1: Structure of a public examination

3. Theoretical Framework: the notion of a public examination market place

The theoretical framework utilised in the analysis in this paper is broadly sociological and follows Cooper (1985) in attempting to "set out the social relationships and interests whose interaction produced change in what counted as school mathematics" (Cooper; 1985, p. 6). As such, it is worth noting that in some countries, although there may be a choice between different qualifications to be made, there is a single syllabus which delivers a qualification. However, as already outlined, this is not the case in England. Currently, schools in England have a choice of six alternative syllabuses which lead to a GCE A-level Mathematics qualification. The advantage of this is to allow to schools a degree of variety in the choice of content and scheme of assessment, although most of the pure mathematical content and all the assessment objectives are currently prescribed for all syllabuses by the Qualifications and Curriculum Authority (QCA), a central regulating agency acting on behalf of the national government. The disadvantage is the requirement to ensure that the standards set for awarding grades are the same for each syllabus.

This leads to the identification of the stakeholders in this 'examination market place' operating in England. These are outlined below.

Students (and their parents) require that the qualification will open doors to higher courses in education, and a future career and employment. The *status* of the qualification is therefore paramount. They require the *assessment* to be *fair*: examination questions should not attempt to be ‘clever’ or ‘original’, but should be predictable and routine. Some prefer alternative forms of assessment to timed examinations, others do not.

Teachers require qualifications which they feel confident will confer success to their students, as their success in the classroom will be measured by this. They therefore look to examining bodies to provide comprehensive guidance on the syllabus and its assessment. They require textbook material which directly matches the demands of examination questions. Some heads of departments will look for specifications which match their own preferences and interests in mathematical content and style – for example, in the choice of applied mathematics offered, or the style of assessment. Past papers and mark schemes, which define the nature of the assessment, are vital to ensure that students are prepared for the examination: as with their students, a ‘good’ examination paper is one which has few surprises, and contains familiar questions which they are confident they have prepared their students to answer successfully.

Examining bodies require that qualifications are *commercially successful*: they must attract sufficient candidates to enable the costs of administration to be covered, and in practice economies of scale mean that the larger the number of candidates the better.

In order to attract candidates, examining bodies need to ensure that the *status* of the qualification is maintained. If examinations are set, and grades awarded, which are perceived to be too harsh on candidates, then schools will seek an alternative which appears to provide their students with a better chance of success. On the other hand, if examinations are made too easy and grades awarded too leniently, then this may devalue the status of the qualification. Moreover, as has already been noted, examining bodies are regulated by the QCA, a government agency, who attempt to ensure comparability of standards between A-level qualifications.

Universities rely upon the qualification to maintain a sufficient supply of qualified students to meet the matriculation standards required to progress to university degree courses. *Mathematics departments* will wish students to have acquired sufficient pure and/or applied mathematical skills to enable them to progress to higher level degree courses, so they have a particular interested in the syllabus specification, and place importance on particular aspects of the construct, such as algebraic manipulation skills, and a thorough understanding of proof. Their concerns, in the face of a proliferation of syllabuses with wide differences in content, led in 1983 to the development of a common core of pure mathematics which is required of all A-level Mathematics qualifications.

Other university departments may require, or prefer, students to have specific applied mathematical knowledge and skills, for mechanics for engineering, statistics for psychology, sound higher level numeracy skills for accountancy, and so on. However, they are less likely to take a detailed interest in the content of an A-level Mathematics qualification, and more likely to require an adequate supply of mathematically qualified undergraduates. They may be less interested in pure mathematical skills such as analytical and algebraic manipulation, and prefer students to develop an understanding of real-world mathematical modelling.

Industry and Commerce have a once-removed interest in the supply of A-level Mathematics qualified students. Although the number of such students who are going directly into employment has declined in recent years, as the graduate employment market has expanded, certain sectors of industry and commerce, such as banking and engineering, for example, have a direct interest in recruiting students with high level mathematical skills, and therefore have a vested interest in maintaining a pool of such potential recruits of an appropriate ‘standard’. Employers’ bodies have a strong political influence on government policy towards mathematical education.

Finally, *national government*, through its agencies, is responsible for overall education policy, and controlling and regulating the examining bodies. A major concern has been to argue for the maintenance of standards – A-level is often referred to in the UK as ‘the gold standard’ – whilst at the same time being able to claim an improvement in educational standards through their policies. Governments have to respond to pressures for reform and change from other stakeholders, such as commerce and industry, or universities. Government policy over the period of existence of A-level has had two over-riding issues which have dominated its direction – the need to increase the size of the graduate workforce, and expand higher education, and the call to end the over-specialisation of the A-level qualification system, in which, until recently, most students studied three or less subjects post-16.

Figure 2 illustrates the pressures on the qualification from stakeholders and their principle concerns. The conflicting interests and agendas of these stake-holding groups create a dynamic for change to the syllabuses servicing the qualification. The 'high stakes' nature of the public examining process ensures that issues such as 'standards', 'reliability', 'fitness for purpose' and 'quality' are frequent matters for debate. For example, the publication of GCE A-level results in the summer dominates the press headlines: the steady increase over the last ten years in the percentage of high grades and passes awarded (referred to as 'grade inflation' has led to annual accusations of lowering of standards. No aspect of government education policy achieves a higher profile in public debate than public examination standards.

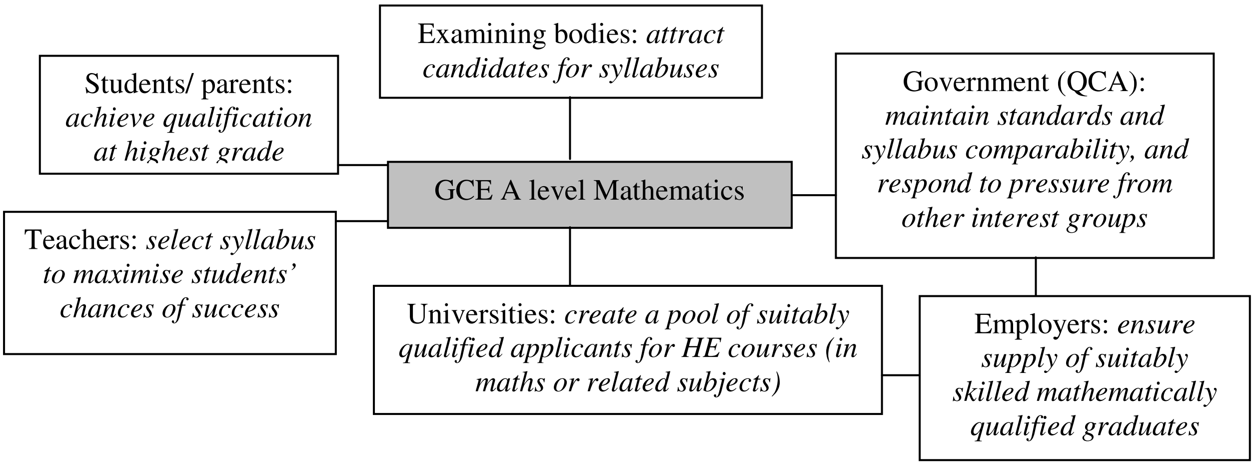

Figure 2: A qualification and its stakeholders

4. Identifying the catalysts for change

In the space available in this paper, a number of key catalysts for change can be highlighted.

(a) Curriculum mismatch between stages of education

The GCE A-level in England is a qualification designed to be taken after two years of study from ages 16 to 18. It therefore needs to make assumptions about the mathematical attainment of students on entry. Changes to the pre-16 curriculum can therefore create a mismatch. This occurred in England with the introduction in 1988 of the General Certificate of Secondary Education (GCSE), replacing GCE Ordinary level and Certificate of Secondary Education (CSE). The effect of this change, without going into detail, was to give access to GCE A-level to students with a far wider range of mathematical experience and attainment. In particular, teachers of A-level could no longer assume that their students arrived on courses with confident and accurate skills in, for example, solving quadratic equations. This curriculum mismatch forced the examining bodies, through pressure from teachers, to reduce the level of algebraic manipulation required in A-level questions. Curriculum materials and textbooks from projects such as SMP 16-19 (Dolan, 1994), which flourished in the 1990s, made fewer assumptions about the attainment of students, as can readily be seen by comparing a popular pre-1990s text such as Bostock and Chandler (1974) with SMP 16-19 Methods (School Mathematics Project, 1991).

This is an example of curriculum mismatch on entry. However, the mismatch can equally occur between school leaving and higher education. In the 1960s, the perceived discontinuity between school mathematics, which was based exclusively on pre-19th century mathematics, and modern university mathematics courses, led to the development of 'modern mathematics' courses, in particular the influential SMP A-level course. For a detailed, sociological analysis of the roots of this change, see Cooper (1985). Some (but not all) of the changes to both the pure and applied mathematics content of A-level proposed by this 1960s project syllabus have gradually been assimilated into all post 1990 syllabuses, which include the language of functions, vectors, transformations, and a greatly expanded, albeit optional, component of statistics.

The stakeholders acting as the catalysers here were leading teachers from the next phase of education, namely lecturers in higher education, together with some leading industrial mathematicians.

(b) Mismatch between supply and demand

Another catalyst for the 1960s reform was the perceived need for training more graduates in mathematics. Prior to the 1960s, demand for places on university mathematics courses exceeded supply, and this mismatch allowed universities to pick and choose from the best of post A-level students. A campaign led by the initiators of SMP with the strong backing of industrialists highlighted the national necessity to train more graduate mathematicians in order to fuel the technological revolution of the 1960s (Cooper, 1985). A-level Mathematics needed to respond by making the courses more accessible, exciting and relevant.

In the 1990s, the qualification was required to respond the expansion of higher education as a whole, following the recommendations of the Robbins Report of 1963.

In both these cases, the demand for producing maths-qualified school leavers outstripped the supply. The assessment 'tap' had to be adjusted so as to turn up the flow. How, or indeed whether, this can be achieved without lowering perceived standards, will be discussed in the next section.

(c) Technological change

New technological advances not only require more mathematically qualified school leavers, but can change the nature of school mathematics. At the onset of our qualification, in 1951 the only calculating aids to students were tables of logarithms. Twenty years later, slide rules had come and gone, and hand-held electronic calculators had become available. In the 1990s, graphics calculators and PCs running graph plotters and spreadsheets had arrived. The issue of how to incorporate these new technologies has always been a major concern of curriculum developers. Indeed, the first ICMI study (Howson and Kahane, 1986) addressed precisely this issue.

Which of our stakeholders carries the flag for technological innovation and initiates syllabus change? Clearly it is the teachers and the school curriculum which places the demands on syllabuses, and in particular schemes of assessment, to adapt and change in the face of calculator and computer technology. However, the need to embrace these technologies has presented dilemmas to the examining bodies, as over-reliance on calculators can allow other useful, albeit more traditional, numeracy and geometrical skills to atrophy. Policy on calculator usage in examinations has fluctuated from embracing it fully to banning. Currently, all A-level schemes of assessment are required by QCA to include non-calculator examination papers.

(d) Changes to the national qualification framework

The above catalysts for change stem from specifically mathematical issues. Change can also be catalysed by more general reform of the qualifications framework. Since its initiation in place of the Higher School Leaving Certificate, GCE A-levels have been criticised for being too narrow a qualification, allowing students to choose freely to specialise in only three subjects. This has led to consistent pressure to reform A-levels in such a way as to broaden the post-16 curriculum. In 1987, the Advanced Supplementary Examination was introduced, equal in standard to the A-level, but equivalent to half an A-level. In 2000, A-level was reformed into a modular qualification, with each A-level syllabus being subdivided into six units, the first three of which were certificated as an Advanced Subsidiary level.

5. The impact of stability and change on examination standards

All the identified stake-holder groups would likely agree on the importance of maintaining the 'standard' of the GCE A-level qualification. Students, parents and teachers have invested in the status conferred by achieving it. Examining bodies need to maintain the quality of their syllabuses compared to competitors. Universities need reassurance that applicants are suitably qualified. Industry and commerce also need to understand and trust the validity of the qualification in measuring the attainment of prospective employees. Governments need to use results as a measure of the success of their policies. It is therefore perhaps not surprising that the maintenance of academic standards is politically crucial, and QCA regularly reports upon subject standards, by using panels of experienced examiners to assess the demand of syllabuses and assessment schemes, and scrutinise the quality of candidates' work at grade boundaries.

In recent years, the procedure by which grades are awarded by the examining bodies has been standardised through a mandatory code of practice, and examining bodies are charged with the maintenance of standards from year to year, using evidence of current and archived candidates' work for A-level is assessed through comparison of performance at GCSE, analysis of centre type, consideration of predicted grades from teachers, and paired paper analysis. Nevertheless, and notwithstanding the increasing sophistication in the use of such tools, judgements of examination standards are subjective: there is no objective measure of how 'easy' or 'hard' a question, or collections of questions, are. Indeed, judgements of the standard of individual examination papers need to take account of historical factors, such as the stability of the syllabus and assessment scheme.

When A-level mathematics was introduced, results were predominantly norm-referenced, with nationally agreed guidelines of 10% grade A, 35% grades A-C, and 70% pass (see Kingdon, 1985, p 74). This quota system of grading was abandoned in 1987 in favour of a greater element of criterion referencing. In 2007, 43.5% of candidates were awarded A grades, 80% grades A-C and 97% passes. Although grades awarded are not a direct measure of standards, this 'grade inflation' has led to many to question whether standards have declined. We shall argue that a gradual increase in the percentages of high grades and passes awarded may be a natural consequence of an effect which we shall call 'case law'.

In a period of stable syllabus, assessment methods and quality of candidature, one would expect that the percentage of candidates awarded grades should remain constant. However, it is possible that incremental changes in grade percentages can, over a period of years, produced pronounced changes, as Figure 3 illustrates.

A Level Mathematics 1992-2001
Percentage achieving each grade

Year	A	B	C	D	E	N	U
1992	20.1	14.6	14.2	14.0	13.0	10.0	14.1
1993	24.3	17.0	15.7	14.0	11.5	7.8	9.7
1994	25.3	18.0	16.5	14.3	11.0	7.3	7.6
1995	27.4	18.3	16.6	14.1	10.5	6.7	6.4
1996	27.5	19.5	17.0	14.1	10.9	5.3	5.6
1997	27.8	20.3	17.3	14.1	10.4	5.1	5.0
1998	29.1	19.5	16.8	13.8	10.6	5.2	5.0
1999	29.2	19.3	16.8	13.8	10.7	5.3	4.9
2000	29.8	19.5	16.8	14.1	10.1	5.1	4.6
2001	30.2	19.0	16.2	14.0	10.7	5.3	4.5

Figure 3: A-level mathematics grades 1992-2001 (Source: QCA)

Does the steady increase in the pass rate and percentage of high grades imply a reduction in standards? One could argue that in a period of stability of candidature, syllabus and assessment scheme, grades should indeed improve, as teachers become more adept at 'teaching to the test' and training candidates to answer questions which can become more and more routine and predictable as 'case law' is established by past papers. For example, a problem of integration such as $\int \sin^2 x dx$ requires the candidate to transform the integral using the double angle formula. The insight required to do is diminished to a learned outcome if the problem is posed and solved regularly from past examination papers. This 'case law' effect is exacerbated if the mathematical content of the syllabus is tightly specified, thus reducing the field of questions which can be tested in the examination.

The 'case law' effect applies equally to other forms of assessment such as coursework, if the tasks set become ossified, and teachers become increasingly expert at training their students to produce responses which qualify for high marks.

Clearly, the predicted rise in 'standards' cannot proceed indefinitely, or eventually all candidates will achieve 'A' grades. The effect is only marginal, and will proceed gradually year by year. It evidently does not apply when norm-referencing is in force, as before 1987, and the percentages of grades awarded is determined independently of the quality of the candidates' work. However, as soon as an element of criterion-referencing is admitted, the 'case law' effect would predict an element of grade inflation, as the criteria by which grades are assessed become more predictable. Does this mean that the mathematical attainment of candidates must perforce continue to rise? Clearly not, since the construct validity (Wiliam, 2007) of the assessment becomes less.

In practice, we have seen that over time syllabuses and assessment methods are revised, and some of the catalysts for change have been outlined in the previous section. The effect of syllabus and assessment change is to undermine to some degree established 'case law' and introduce new, less predictable, elements into the assessment. It is to be expected that this will lower the technical competence required by candidates, especially on new mathematical content which may be less expertly taught in the classroom. Even though performance on novel examination questions would be expected to decline, it would be unfair to candidates for this to be reflected in the percentages of grades awarded: these will be maintained notwithstanding.

The above analysis would imply a pattern of grade awarding as pictured in Figure 4.

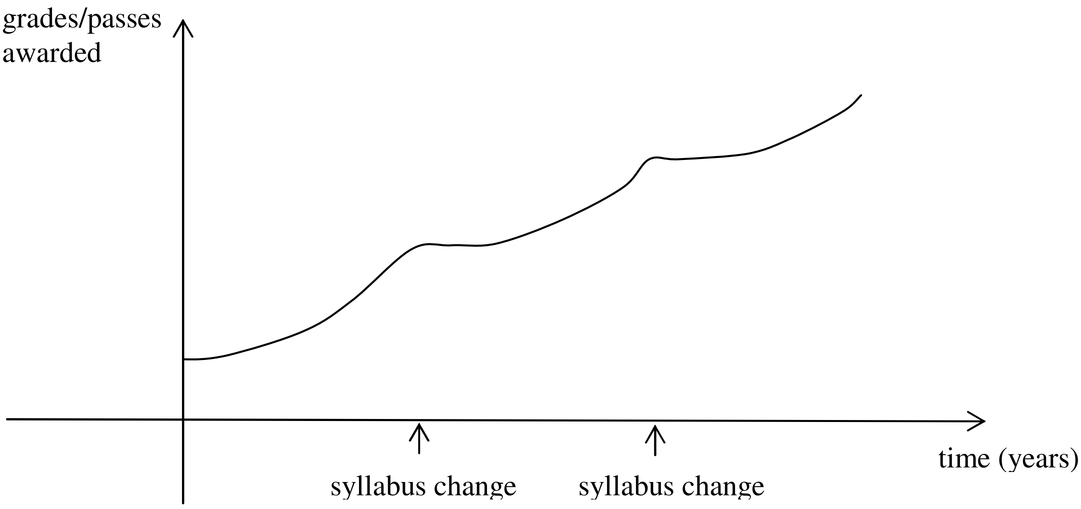

Figure 4: the 'case law' effect on grade awarding

In practice, the above model is too simple to explain longitudinal patterns in the awarding of grades. The extent to which the 'case law' effect distorts examinations as a measure of mathematical attainment depends upon other factors, such as the breadth of the syllabus content, and the variety of the questions and assessment methods adopted by the syllabus. Here, there is often a tension between validity and reliability of test items. Indeed, it can be argued (Little, 1993) that the reliability and validity of assessment tools are in inverse proportion. A question such as the integral example discussed earlier, which is well established by case law, tests elements of the construct reliably, and is accepted as 'fair' by students and teachers; but if it becomes a 'trick', a purely learned response, it loses validity as a test of understanding. On the other hand, more open-ended questions, which demand more imagination from candidates, may appear more valid tests of mathematical attainment, but at the expense of reliability. An extreme example of imagination getting the better of reliability is the following SMP A-level question.

'As the sun was setting in a clear African sky, it was noticed in a Super VC10 flying north that the outline of the westward windows was projected on the other side of the cabin about 6 inches above the windows on that side. Estimate roughly the height of the aircraft. (SMP A-level Paper 2, July 1968)'.

This question provides ample scope for classroom discussion, but as a five minute question from a timed written paper, it is an inappropriate test which only the ablest candidates would know how to start.

Placing mathematical questions in non-mathematical contexts has been criticised on the grounds of artificiality (Wiliam, 1997), realism (Boaler, 1994) and reliability – see, for example, Cooper and Dunne (2000), Ahmed and Pollitt (2007). Introducing a context automatically adds to their wordiness, and comprehension may be regarded as a construct-irrelevant variant (Wiliam, 2007). However, they do demand students to apply their mathematical understanding in less familiar situations, and are less subject to the 'case law' effect. If mathematical modelling is admitted into the examination construct, then a wider, less routine, palate of question style becomes admissible, and coursework a more valid assessment tool than timed written papers.

An example of an assessment tool which clearly avoids the problem of 'routinisation' of examination questions is the *comprehension paper*, in which students are presented with an article which explains a mathematical model of a real-life situation, and required to answer questions which test their understanding of the article and the model. This is currently used in one A-level syllabus (Oxford Cambridge and RSA Examinations, 2004). The 'backwash' effect of such an assessment tool, in encouraging students to read mathematical articles and understand how mathematics can indeed be a useful tool in real life, is highly beneficial; the 'case law' effect is negligible; but the comprehension skills required may be said to reduce the reliability of the tool, for example in assessing students whose first language is not English.

6 The assessment of 'Coursework'

An analysis of the influences on GCE A-level mathematics would be incomplete without addressing the issue of coursework assessment. Key questions concern what caused the rise in popularity in these alternative forms of assessment in the 1990s, and what has caused their decline in the 2000s. The educational rationale for coursework assessment was established in the influential Cockcroft report (Cockcroft, 1982), which insisted on the importance of assessing practical and investigational work in the GCSE Mathematics construct. This led directly to the development of project syllabuses at A-level which included elements of coursework, including one syllabus which included 100% coursework[3].

In order to sustain enthusiasm for such initiatives over time, stakeholders need to be convinced of the benefits. Students are divided on the merits of coursework, although research, for example Goulding (1995), Little (2007), tends to suggest that girls perform better, and prefer, coursework, to boys. Teachers also are ambivalent. Whilst some enjoy the control they can exercise on part of the assessment, and see the educational benefits, most are less positive about the time taken to assess pieces, and their ability to assess the work reliably. Universities, industry and commerce are looking for candidates to demonstrate many of the qualities required by coursework, such as research and enquiry skills, but are reluctant for curriculum time to be taken away from the traditional content–based learning of conventional A-level classrooms.

Perhaps the main catalyst for the decline in coursework in England has been the increasing usage made of public examination results to measure the output of the educational system. Since 1994, GCSE and A-level results have been used to compile league tables of the performance of schools. Thus this qualification became a measure not just of the attainment of students, but of teachers and their schools. Under these circumstances, it becomes undesirable for teachers to have control over the tools through which their performance will be assessed.

7 Conclusion

This paper describes the infrastructure of a particular qualification (identified a stakeholder model), proposes some catalysts for change in the syllabus, and discusses some issue relating to syllabus content, standards, validity and reliability. It is hoped that this might be useful when considering other systems of qualification across the world.

[3] The 'RAMP' (Raising Achievement in Mathematics) project, developed by West Sussex Institute of Higher Education.

The most interesting feature of the qualification framework as it has developed in England is the 'market' created by having multiple examining agencies delivering a single qualification. Whilst this does create problems of comparability, and can lead to pressures which are not wholly educationally desirable, the market does create a dynamic for change which a public examination monopoly might not readily achieve.

Acknowledgements

We gratefully acknowledge the support of the *London Mathematical Society* and the *University of Southampton* in providing bursaries to enable a version of this paper to be presented at the *11th International Congress on Mathematical Education* (ICME11), Monterrey, Mexico, July 2008.

References

AHMED, A. & POLLITT, A. (2007) Improving the quality of contextualised questions: an experimental investigation of focus. *Assessment in Education: Principles, Policy and Practice,* 14(2)**,** 201-232.

BOALER, J. (1994) When do girls prefer football to fashion? An analysis of female under-achievement in relation to realistic mathematics contexts. *British Educational Research Journal,* 20**,** 551-564.

COCKCROFT, W. H. (1982) *Mathematics Counts,* London, HMSO.

COOPER, B. (1985) *Renegotiating Secondary School Mathematics,* London, Falmer.

COOPER, B. & DUNNE, M. (2000) *Assessing Children's Mathematical Knowledge,* Buckingham, Open University Press.

DOLAN, S. (1994) 16-19 Mathematics. *Teaching Mathematics and Its Applications,* 13(1), 28-33.

GOULDING, M. (1995) GCSE Coursework in Mathematics: Teachers' Perspectives and the Performance of Girls. *Evaluation and Research in Education,* 9(3)**,** 111-119.

GRIFFITHS, H. B. & HOWSON, A. G. (1974) *Mathematics: Society and Curricula,* Cambridge, CUP.

HOWSON, A. G. & KAHANE, J. P. (1986) *The Influence of Computers and Informatics on Mathematics and its Teaching,* Cambridge, CUP.

LITTLE, C. (1993) The School Mathematics Project: Some Secondary School Assessment Initiatives in England. In NISS, M. (Ed.) *Case Studies in Assessment in Mathematics Education.* Dordrecht, Kluwer.

LITTLE, C. (2007) A Coursework Task in A-level Mathematics: a Survey of Student opinion. *British Society for Research into Learning Mathematics,* 27(3)**,** 78-83.

NISS, M. (1993a) *Case Studies in Assessment in Mathematics Education,* Dordrecht, Kluwer

NISS, M. (1993b) *Investigations into Assessment in Mathematics Education,* Dordrecht, Kluwer.

OXFORD CAMBRIDGE AND RSA EXAMINATIONS (2004) *MEI Structured Mathematics Specification*, Cambridge, OCR.

SCHOOL MATHEMATICS PROJECT (1991) *SMP 16-19 Methods,* Cambridge, CUP.

WILIAM, D. (1997) Relevance as MacGuffin in mathematics education. Paper presented at the *1997 British Educational Research Association Conference*, York

WILIAM, D. (2007) Quality in Assessment. In SWAFFIELD, S. (Ed.) *Unlocking Assessment.* London, David Fulton.