

Examination Reform: Impact of Linear and Modular Examinations at GCSE

Summary Report

Jo-Anne Baird, Daniel Caro, Victoria Elliott, Yasmine El Masri, Jenni Ingram, Tina Isaacs, Anne Pinot de Moira, Ashmita Randhawa, Gordon Stobart, Michelle Meadows, Caroline Morin, and Rachel Taylor

OUCEA/19/1
Ofqual/19/6506/2

April 2019

Foreword by the Chief Regulator

GCSE examinations are highly respected the world over. Nonetheless, what students learn in secondary school needs to keep pace with changes in society beyond the school gate to ensure that young people are working towards the right qualifications to prepare them for life and work. Any change should be conducted in a well-planned manner. Further, it is important that we learn from the changes we make so that we can monitor what is effective and what might need to be adapted in terms of the content of the examination reforms or the process of reforming examinations. This project is part of Ofqual's work to ensure that examination reforms are operating well for the young people who take them.

In 2011 the vast majority of GCSEs were modular examinations. From 2012, the rules around GCSEs were modified to require students to take all of the assessments at the end of the course. This change turned GCSEs designed to be modular into linear qualifications. The more fundamental reforms begun in 2013 cemented this move to linearity, redesigning the syllabuses and examination structure to fully deliver linear qualifications. Reforming the examinations in this way was a response to concerns that modular examinations led to constant testing and were partly responsible for a perception that examination standards had declined in England. To ensure that the change was manageable, the reforms were phased over a number of years. Schools, colleges and teachers have had to respond to a considerable period of reform.

Existing research evidence regarding the impact of examination structure on teaching and learning is surprisingly sparse. Understanding how the change of structure of the examinations has affected examination standards, teachers' reported classroom practices, who is entered for the examinations, the costs of GCSEs and teachers' views on how the changes have affected students is clearly important in ensuring that the reforms are operating as intended. Hence, this research makes an important contribution to Ofqual's evaluation of the reformed GCSEs. Conducting research on the effects of examination policy is usually a lengthy process because it takes some time to make the changes, put them into place in schools and colleges and then for the effects of the first examinations to be felt. This is a three-year project in which Ofqual has collaborated with Oxford University's Centre for Educational Assessment.

The reform to the examination structure did not happen in isolation. There were a number of concurrent policy changes. Early and multiple entries for GCSEs were discouraged, Progress 8 was introduced in school performance table calculations, the curriculum was updated and made more demanding, and the 9 to 1 grading replaced the previous letter grading of GCSEs. At the same time, school structures have continued to evolve, there have been teacher recruitment and retention drives and school budgets have been under discussion. All of these factors will have had their own impact upon school and classroom practices.

Examination Reform: Impact of Linear and Modular Examinations at GCSE

This report has important messages about the effects of structural reform upon grading outcomes, teachers' changes of practice and the impact of the reform for different groups of pupils. We conclude from the range of evidence gathered, that in the current educational context, linear examinations are more suitable at GCSE than modular. We continue to monitor the effects of policy reforms upon the education system.

Sally Collier
April 2019

Contents

What is meant by ‘modular’ and ‘linear’ GCSE examinations?	4
GCSE Examination Structural Reform initiated in 2013.....	4
Do other national examination systems have modular examinations?.....	5
Did modular examinations affect grading standards?.....	6
Are modular or linear exams more equitable?	8
What were the economic impacts of the structure change?	8
Effects of the reform on school practices and students	9
Conclusions.....	10
References.....	11

What is meant by ‘modular’ and ‘linear’ GCSE examinations?

A modular examination is one in which the totality of the assessment is broken into discrete units for assessment, the results of which are combined to give an overall result. Various forms of modular examinations exist. GCSEs were mainly of the internal form (Ertl and Hayward, 2010), in which the qualifications were subdivided into units that were particular to that qualification, credit for modules could not be carried across qualifications and there was a separate syllabus for each qualification. A common feature of modular assessments is the facility to re-sit modular examinations. Later policy restrictions for the modular GCSE examinations meant that students could re-sit each module once and 40% of the examination could only be taken at the end of the course. Linear qualifications require all examinations to be taken at the end of the course of study.

GCSE Examination Structural Reform initiated in 2013

GCSE examinations were introduced for first examination in 1988. They replaced the Ordinary Level (O Level) and Certificate of Secondary Examination (CSE) qualifications, with a single examination designed to suit the entire cohort of 16-year-olds. Reforms of the examinations have been conducted periodically, either as a suite of examinations, or for particular subjects, with a rolling programme of changes being implemented. Fifty-three GCSE subject areas are currently available and the highest entries are for mathematics and English.¹ In 2018, there were over 5.2 million GCSE results issued. An average of eight subjects was taken by pupils, but this varied and almost a quarter of students took ten or more subjects.

Tiered examinations have been available in a number of subjects since GCSEs were introduced, with the higher tier examination being designed for students who were expected to gain at least a grade C and the foundation tier designed for students who were not expected to get a grade C or higher. A further set of reforms begun in 2013 reduced the number of subjects with tiered structures and where they were continued, the higher tier was designed to be suitable for students who were expected to gain at least a grade 5.

At the time of their introduction, GCSEs were the school leaving examination but since 2015 young people must stay in education or training until they are 18 years old. The GCSE examination has largely been an end of course assessment for most of its history. Some modular courses were introduced, such as in science, but these were in the minority of syllabuses available in most years. Reforms of the GCSE were initiated in 2007 for examinations in 2011. New examinations for GCSE English, mathematics and ICT came later, in 2012. These saw all but GCSE mathematics redesigned such that the examinations were

¹ <https://www.gov.uk/government/publications/get-the-facts-gcse-and-a-level-reform/get-the-facts-gcse-reform>

Examination Reform: Impact of Linear and Modular Examinations at GCSE

modular. Introduction of modular GCSEs in the 2007 reforms can be explained by two factors; commercial drivers and policy facilitation.

Examination boards proposed modular GCSEs because they were popular with schools. Since their competitors (the other examining boards) were producing modular proposals, it was difficult for a single board to take a different stance for commercial reasons. To offer linear examinations in large entry subjects in this market context could have decimated the market share of an exam board. Policy could have acted as a barrier to the wholesale introduction of modular examinations but at that time policy-makers at the Qualifications and Curriculum Authority (QCA) supported the production of assessments with modules that would be interchangeable across qualifications. This was in keeping with international policy developments in higher education (The Bologna Framework)² and vocational education (qualifications frameworks such as the Qualifications Credit Framework (QCF)).

In 2013, the then Secretary of State for Education, Michael Gove, introduced further reforms to tackle a number of concerns about the examination system and the standards of examinations, which included a return to end of course examinations. The demand of GCSEs was increased by changing the subject content of the syllabuses (specifications) and the style of the examinations with, for example, more open ended, essay questions being introduced. The Government aimed to increase the level of challenge of the GCSE and thereby better prepare pupils for work and further study. Coursework (controlled assessment) was also reduced because of concerns about formulaic teaching and learning and a lack of reliability, but it is still a feature of subjects in which performance or production of physical artifacts are an integral part; such as drama, physical education, art and so on. The structure of the qualifications was changed from modular to linear so as to reduce the disruption to teaching and learning caused by repeated formal assessment.

Do other national examination systems have modular examinations?

When evaluating the change from modular to linear examinations, it is useful to ask whether the GCSEs were unusual in their modular structure. This project investigated the summative assessments of 18 jurisdictions plus the International Baccalaureate at lower and upper secondary levels of the education systems; Australia (New South Wales, Queensland), Canada (Alberta, Ontario), China (Hong Kong, Shanghai), Finland, France, Germany, Ireland, Japan, Korea, the Netherlands, New Zealand, Scotland, Singapore and the US (Florida, Massachusetts).

² <http://www.eua.be/eua-work-and-policy-area/building-the-european-higher-education-area/bologna-basics/Bologna-an-overview-of-the-main-elements.aspx>

Examination Reform: Impact of Linear and Modular Examinations at GCSE

By far the most common assessment structure was linear examinations. Exceptions to this, where modular examinations were in place, were France (lower secondary), Queensland (upper secondary) and Scotland (upper and lower secondary). Therefore, although modular examinations were unusual, they were to be found in some systems. Further, most of the systems had elements of staged assessments because teacher-assessments were often incorporated into students' grades or in some cases determined them altogether.

Modular assessments have been commonplace in higher education for decades, with entirely end of course assessment of degrees now being a rarity. In the UK, A levels were reformed into modular examinations for a period and Higher examinations in Scotland became modular in the early part of this century. Vocational assessments have a long history of using modular examination structures so that students' courses can combine different areas of study that best suit their occupational needs. At the time of the blanket introduction of modular GCSEs, qualifications frameworks were being revised or introduced in the UK and vocational, Diploma courses were being heavily emphasised by policy-makers. Although relatively unusual for secondary examination systems around the world, modular GCSEs were seen at the time as a logical step in the integration of academic and vocational qualifications in a broader framework. Interchangeability and personalised curricula are seen as highly desirable features of such systems.

Did modular examinations affect grading standards?

Concerns about the effect of modular examinations upon examination standards were part of the driver for the 2013 reforms. To investigate whether examination structure affected outcomes at GCSE, we analysed national data collated at syllabus level (2002 – 2014) and the national pupil database (2007 – 2014). Our research focused upon outcomes in GCSE English, mathematics and science. We found no educationally significant evidence that GCSE outcomes were affected by the structure of the examinations. Further, grades awarded at A level were not statistically significantly affected by whether students had sat modular or linear GCSEs.

Our analyses compared students who sat modular examinations with those who sat linear examinations but in any one year examination structures were usually common to any one GCSE subject. Having a choice between modular and linear examination syllabuses within a year was unusual. This means that separating the effects of examination structure changes from other national changes between years was not possible. For example, subject content changed at the same time as examination structures changed which could have resulted in more demanding assessments. It is impossible to disentangle these effects.

Much research has been conducted on pupils' re-sitting patterns and re-sit effects on outcomes in the course of sitting GCSEs. Most students who re-sat GCSE mathematics or English did better than on their first sitting, with re-sitting being far more prevalent in mathematics than in English (Vidal Rodiero and

Examination Reform: Impact of Linear and Modular Examinations at GCSE

Nádas, 2012). Due to the wealth of previous research on re-sitting, our analyses focused upon the overall examination grades awarded, not the effects of module re-sits.

A number of other policies related to examination entry influenced results during the same period of time in which the reforms to modular and linear examinations took place. School Performance Tables incentivised schools to ensure that students gained at least a grade C in their GCSEs. Before the reforms, schools had begun to enter students for GCSEs earlier (students were younger), more frequently than in the past, multiple times for the same GCSE qualification and entered them for the same qualification with multiple examination boards. Thus, School Performance Table pressures gave rise to these unintended consequences and changes were subsequently made to the measures such as the Progress 8 measure,³ that were designed to reduce these undesirable behaviours.

Setting of examination standards was made more difficult by the introduction of modular structures because students were able to put a lot of credit into the examination bank of results prior to the overall standard setting for the GCSE being conducted. This meant that examination boards did not have as much leverage over the final outcomes; an issue that was discovered in 2002 when the modular A level examinations were first issued and in 2012 when the first modular GCSE English results were released. Early entries also made the standard setting process problematical because it was difficult to use the same system for gauging how the cohort of students was likely to perform on the examination when a large proportion of them had taken the examination without the same level of maturity and possibly even curriculum exposure and study as in the past.

Ofqual's comparable outcomes approach, which is applied by all of the examination boards, was introduced to GCSE grading in 2011 (Ofqual, 2015). This approach was designed to ensure comparability both between examination boards and between years in the same subject. Although there were challenges in applying the system with the entry policies outlined above, it was clearly robust to these challenges in terms of any differential outcome standards in modular or linear examinations.

A number of changes to the examination system were enacted at national level concurrently that affected both schools' entry policies and examination outcomes. The raft of reforms were designed at least in part to tackle concerns about examination standards. Changing the structure from modular to linear examinations did not, in itself, halt rises in examination results. The comparable outcomes approach tackled grade inflation. However, structural reform to GCSE begun in 2013 played a small part in reducing year on year rises in examination results because the comparable outcomes procedures could be applied without added complications for linear examinations.

³https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/676184/Secondary_accountability_measures_January_2018.pdf

Are modular or linear exams more equitable?

Analyses of the effects of examination structure for gaps in outcome by school type, gender, socioeconomic status and progression to A level were conducted for results in GCSE English, mathematics and science. It is important to note that findings may simply reflect schools' policies with regard to the suitability of examination route, in years where choice was offered. Students were not randomly allocated to the different examination structures – quite the opposite.

That said, analyses taking into account prior attainment did not show educationally significant effects of examination route in terms of changing the attainment gap for gender or socioeconomic status. Outcomes at A level were not affected to an educationally significant extent by prior entry for modular or linear GCSE examinations either. Overall there was no statistical evidence to suggest a difference between modular or linear examinations for equity in examination outcomes by school type, gender, socioeconomic status or progression to A level.

Nonetheless, the national pupil database provides a wealth of data for analysis and despite the extensive research presented here, it is possible that the introduction of other variables or interaction terms and the application of different statistical techniques could provide more information on the equity effects of examination structure. Exploration of the national pupil database using a wider range of models could be a fruitful avenue for future research.⁴

What were the economic impacts of the structure change?

England's examination market is a regulated oligopoly (Jones, 2011), with a few organisations dominating the market. Examination boards use an 'invest and harvest' economic strategy, as the costs of reform are high and need to be recouped over time. Additionally, many GCSE subjects are not profit-making, so there is a great deal of cross-subsidy within a year. Indeed cross-subsidy often goes across the suite of qualifications offered by a board.

This makes using pre-existing data to assess the cost of examination reform and the impact on examination fees difficult. Nonetheless, between 2015 and 2018 price rises were in general in excess of the retail price index, coinciding with a period of reform in which the examination boards would have to invest.

Exam fee costs to schools and colleges have also been affected by a number of structural changes, including module re-sits and multiple entries. In 2013, double the number of pupils in the cohort was entered for GCSE mathematics

⁴ Analyses taking into account prior attainment could not be conducted for comparisons of independent school results because Key Stage 2 examination data is not available for those schools.

Examination Reform: Impact of Linear and Modular Examinations at GCSE

examinations, reflecting a high volume of multiple entries. Late entries, which can cost double the fee of a standard entry, were less frequent in linear examination structures.

All of these factors would reflect costs to school and college budgets. Offset against these costs is the use of teacher time for within-school (mock) assessments, which were reportedly more frequent after the linear structure was introduced. However, staff costs are typically a sunk cost rather than an additional cost in terms of additional assessments. Overall, linear examinations are less costly to the education system, especially in combination with diminished multiple and late entries.

Effects of the reform on school practices and students

The effects of the reform on school practices and students were explored through two sets of interviews conducted with teachers of English, mathematics and head teachers. In total, 84 interviews were conducted and two thirds of those interviewed in the second phase were the same individuals that we spoke with in the baseline research, before the linear syllabuses were introduced.

Prior to the introduction of the linear examinations, teachers were very concerned about the effect of the changes upon examination outcomes. These concerns also related to the increased demands of the subject content that formed part of the policy reform. Many teachers commented that they did not have enough information about the reforms and that the lead-in period was too short.

Following the reforms, a number of teachers said that they considered the linear examinations to be a fairer reflection of students' performances. These comments could be related to the inflationary effects that re-sits have, due to inherent error in any assessment system (see, for example, Wheadon, 2010).

Teachers reported having introduced more examinations within the school and more examination preparation in schools due to the introduction of linear examinations. It seems that frequent modular examinations were useful for learning examination technique as well as for getting feedback on progress. It is possible though that school-produced examinations will reduce once teachers become more experienced and confident in delivering the new qualifications.

Whilst teachers were more comfortable overall with the reforms in the interviews conducted after their introduction, the effects on sub-groups of students were of concern. Importantly, student mental health was cited as an issue. Linear examinations were felt to cause more pressure due to their high stakes, all-or-nothing, performance on the day nature. There is some evidence to suggest that mental health problems are growing in society and for young people in particular. There is, however, no hard evidence to support a cause and effect relationship between student stress and the changes to GCSEs.

Examination Reform: Impact of Linear and Modular Examinations at GCSE

Nonetheless, this wider societal problem may be exacerbated for some young people by the change to examination structure. That said, some pupils may have found the constant pressure of modular examinations more damaging. Indeed, modular examinations were previously criticised for maintaining a constant level of examination pressure throughout the period of study. We do not know whether continuous, but lower stakes pressure was detrimental to a wider range of pupils than the effects of shorter periods of higher stakes pressures associated with linear examinations. It is likely that different pupils will respond to each of these types of pressures in different ways.

Overall, teachers were beginning to adjust to the changed demands of the national curriculum, the reform to examination structure, the new grading scale and so on. However, there was a general feeling in both phases of the qualitative research with teachers that a moratorium on reforms to general qualifications would be welcome. In 2018 the Secretary of State committed to no further reform of GCSEs, beyond the changes already announced, for the rest of this parliament.⁵

Conclusions

Internationally, modular secondary school examinations are not commonplace. Further, it was a permissive environment rather than proactive policy choice which led to the introduction of England's suite of modular GCSEs in 2007. Policy makers may wish to reflect on the contexts in which central control of key qualification design decisions are prudent, and those contexts where the market is best placed to determine design.

We found no educationally significant effect of examination structure on grade outcomes in English, mathematics or science. Nor did we find any statistical evidence of effects of structure upon progression to A level. So there was no evidence to suggest that either modular or linear GCSEs led to better educational outcomes. Importantly, equity gaps in gender, socioeconomic status and school type were not affected by structure once prior attainment was taken into account. While some teachers believed that certain groups of students would perform better in a modular system, this view was not supported by analysis of outcome data, where such beliefs could be investigated.

Once teaching was well underway, many teachers considered that linear GCSEs provided more valid assessments of students' performances than did modular examinations. However, concerns about students' mental health were raised, with linear examinations considered to have had a negative impact upon wellbeing for some students. This suggests that well targeted support for students who suffer from examination anxiety is needed.

⁵ <https://www.gov.uk/government/news/damian-hinds-sets-out-plans-to-help-tackle-teacher-workload>

Examination Reform: Impact of Linear and Modular Examinations at GCSE

This may be especially important as we found that more internal, school examinations and more examination preparation had been introduced as a result of the change from modular to linear GCSEs. It may be that the internal assessments will reduce as teachers become more familiar and confident with the new qualifications. While the treadmill of formal examinations has reduced, the corresponding increase in mock examinations means that assessment may still heavily feature in some students' educational experiences.

Increases in grade outcomes (at the time of reform suspected to be unwarranted by policy makers) was tackled by the introduction by Ofqual of the comparable outcomes methodology. Therefore, the change to examination structure did not impact upon this directly. However, there is evidence that the maintenance of standards is more straightforward for linear GCSEs and there have been instances where fairness and public confidence has been undermined by difficulties in grading modular GCSEs. Where qualifications are high stakes for students and teachers, a linear approach is more likely to facilitate comparable standards over time and between exam boards.

Although linear GCSEs are less expensive to the education system than modular GCSEs, the reforms had a significant cost. Teachers adapted to the changed qualifications in a range of ways; there was a lot of upheaval in the system. Indeed, teachers felt that a moratorium on reforms to general qualifications would be welcome. As such, qualification reform must produce considerable benefits for it to be worthwhile.

Finally, it is important to note that these findings relate only to GCSE. The advantages and disadvantages of modularity versus linearity are likely to vary with the purpose and educational context of any qualification. Better understanding the impact of different design choices on the maintenance of standards, outcomes and teaching and learning is a crucial area of further research.

References

- Ertl, H. and Hayward, G. (2010), Modularization in Vocational Education and Training. In: Penelope Peterson, Eva Baker, Barry McGaw, (Editors), *International Encyclopedia of Education*. Volume 8, pp. 383-390. Oxford: Elsevier.
- Ofqual (2015). *Setting GCSE, AS and A level standards in summer 2014 and 2015*. Retrieved from https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/451321/2015-08-05-summer-series-gcse-as-and-a-level-grade-standards.pdf
- Vidal Rodeiro, C.L., & Nádas, R. (2012). Effects of modularity, certification session and re-sits on examination performance. *Assessment in Education: Principles, Policy & Practice*, 19(4), 411-430, DOI: 10.1080/0969594X.2011.614218.
- Wheaton, C. (2010). *The impact of re-sitting on the fairness of grading*. RPA_10_CBW_RP_023. CERP, Guildford.