Playing the system: incentives to 'game' and educational ethics in school examination entry policies in England

J. Ingram[a]*, V. Elliott[a], C. Morin[b], A. Randhawa[a], and C. Brown[c]

[a]Department of Education, University of Oxford, Oxford, UK; https://orcid.org/0000-0003-4118-2413

[b]*Office of Qualifications and Examinations Regulation, Coventry, United Kingdom;*

[c]*School of Education, Oxford Brookes University, Oxford, UK;*

*corresponding author: Department of Education, University of Oxford, 15 Norham Gardens, Oxford, OX2 6PY, Jenni.Ingram@education.ox.ac.uk, https://orcid.org/0000-0003-4118-2413

Acknowledgements

Playing the system: incentives to 'game', and educational ethics in school examination entry policies in England

There has been a period of intense policy change involving GCSE examinations in England, proposed partly in response to schools using tactics to maximise performance against accountability measures. The reforms included a change to linear rather than modular entry, removing partial re-sits, and limiting early and multiple entry to examinations by changing school accountability measures. We present new empirical data from interviews conducted with senior teachers at 15 schools. The focus of these interviews has been in the English and Mathematics departments; the first subjects to be examined in the new specifications. The data suggest that teachers acknowledge this practice of 'gaming' but only as something 'other' schools did. Whilst the reforms have now allowed for the system to be viewed as a more level playing field, teachers still describe a constant tension in the decisions surrounding examination entry. They describe the desire for a balance that is not just between school and student outcomes, but also between different outcomes such as motivation, performance, and engagement. Tensions arise between these outcomes when entry choices are being made.

Keywords: qualifications; gaming; examination entries; GCSE

Introduction

The qualification system in England has undergone considerable change over the last few years, and this has been accompanied by significant curriculum and policy changes. These changes were in response to perceived issues with the previous system, but were primarily driven by a desire for the exams to be more demanding. Baird et al. (2013) discuss concerns about the rigour of assessments, grade inflation, the use of multiple re-sits and the validity of teacher assessment, among others, as possible justifications for reform. Head teachers, and their staff, have a dual role in relation to their decisions about examinations and qualifications: a duty to their students and a consideration of the needs of their school (Wilson, Croxton & Atkinson, 2006). In this paper we explore the perspectives of teachers both in anticipation of the changes and following the changes in terms of how examination

practices have changed. We consider the ethical and educational tensions, as perceived by teachers, over examination entry policies through this period of change.

The General Certificate of Secondary Education (GCSE), taken by 16-year olds in England in a number of subjects, is a high stakes qualification which was first taken in 1988. In England, the Department for Education is responsible for setting the curriculum as well as establishing which qualifications are included in performance tables. The Office of Qualifications and Examinations Regulation (Ofqual) is responsible for the assessment objectives in each GCSE subject and regulates the assessments and qualifications developed by the different awarding organisations based on the content specified by the Department for Education and schools and teachers have a choice of qualifications in each subject offered by different awarding organisations. Finally, schools are held accountable for the quality of their teaching and students' learning by the Office for Standards in Education, Children's Services and Skills (Ofsted), who is responsible for inspecting and giving ratings to a range of educational institutions.

The qualification system in England for 16-year olds is very high stakes, for both students and schools, and the rules of the system can be interpreted in different ways to achieve different outcomes. Such behaviours, following the wording rather than the spirit of the rules, can be deemed to be 'gaming. A good pass in GCSE English and mathematics is usually essential for employment or further study. The proportion of students gaining a good pass within a school is also published in the national press. Schools are then inspected and judged on these figures which can have a huge impact on the day-to-day running of a school. Playing the system involves decisions that can impact on both student outcomes and school outcomes, but, as the teachers in our study report, may also have wider implications on student motivation and engagement with education.

Accountability measures and their impact

It is a key aphorism in performance management that 'what gets measured gets done' (Behn, 2003: 533), and Wilson, Croxton & Atkinson (2006) concluded that accountability measures in education (school accountability and the inspection system) have unintended consequences as a result. Performance tables (widely known as league tables) were introduced in England in 1992 to make schools accountable to key stakeholders and to provide parents with information about schools' performance. The key measure reported was the percentage of 16-year old pupils achieving 5 or more good passes where a good pass was defined as a grade of C or above (A* to C). As a result, this became a key focus of effort in schools.

In 2004, the UK government introduced a "floor standard" to identify schools failing to meet minimum performance expectations. The "floor standard" is the minimum percentage of pupils in a school that must achieve five good passes including English and mathematics. Schools below the "floor standard" could expect increased monitoring and inspections. Another measure, the English Baccalaureate (EBacc), was introduced in 2010 looking at the percentage of pupils achieving a good pass in English, mathematics, the sciences (including computer science), a language and a humanity subject (history or geography).

As schools are evaluated using these measures, teachers are under pressure to meet these targets and some schools use the measures to assess teachers' performance. This, in turn, leads to schools' and teachers' practices that are not always in the best interest of pupils. De Wolf and Janssens (2007: 382) offer two alternative terms for behaviour designed to maximise performance in response to accountability measures: 'intended strategic behaviours' or 'gaming'. For example, a number of teachers constantly review pupils' performance and focus on pupils who are just below the C grade (NAO, 2003). This can lead to higher attaining pupils not progressing as much as they could because the teacher's time is targeted towards borderline pupils. Such accountability measures can also influence teachers

4

to teach to the test and to not cover the whole curriculum to focus on areas which are more likely to be assessed (Sturman, 2003), which de Wolf and Janssens would consider 'unintended strategic behaviour' (2007: 382). Smith (1993) suggests that strategic intended behaviour in relation to accountability measures can go so far as to be considered fraud, deception or misrepresentation. While few would go so far as to suggest that boosting student grades constituted fraud or deception when it did not involve outright cheating, one of the concerns raised in the examination reforms, specifically the change to linear specifications, which are the background to this study was implicitly that of misrepresentation: that a student's grade did not represent a true indication of their potential attainment at the end of a two-year course.

The key accountability measure at GCSE changed in the summer of 2016 from the five good passes measure to Progress 8, a measure of progress from key stage 2 to key stage 4 in eight subjects, with mathematics (and the better of English language/English literature) counting double (Department for Education, 2014). The change to Progress 8 as the headline accountability measure might dissipate the focus on one specific grade but previous accountability measures continue to be reported and the fact that mathematics and English Language/Literature have a greater weight will mean that the pressure will still be on schools and teachers to improve the achievements of their pupils in those subjects, that is, the ones under consideration in this study.

One of the key ideas explored in the data is the idea of 'gaming' examinations through entry practices. Such behaviour due to increasingly stringent accountability regimes has been explored both internationally (Ravitch, 2010) and nationally (Perryman et al, 2011) where it has been recounted that teachers felt pressure to not only get good results from themselves and their students but also for the whole school. The entry practices which were tackled by the recent examination reform, including but not limited to linearisation, and

therefore raised in the data are considered in the next section but researchers have also

suggested other related behaviours, such as the diagnosis of dyslexia or other specific

learning difficulties driven by middle-class parents to ensure resources (such as extra time in

assessments) for their children who may not otherwise succeed in the examination system

(Beck, 1992; Riddell & Weedon, 2006) or schools assigning students to special educational

provision which is exempt from accountability measures (e.g. Figlio & Getzler, 2006).

The recent reform of General Certificate of Secondary Examinations (GCSEs)

Reformed GCSEs in England were first taken in June 2017, having been taught from

September 2015. Several different features of the examination system which were open to

'perverse incentives' were changed to varying degrees and are considered below.

It has been shown that teachers have used early and multiple examination entries to

maximise the pupils' chances of achieving a grade C (Taylor, 2016). Early entry means that

pupils achieve a qualification before they are 16-years old. Previously, early entries were

used by schools to enter strong candidates who were likely to achieve a good grade but in

later years schools started to enter a wider range of pupils early, with a variety of

justifications for this practice. This change in early entry practice might also be partly due to

the national tests for 14-year olds being abolished in 2009 leaving teachers with the

opportunity to start teaching the GCSE curriculum earlier. Further evidence of this trend

comes from the composition of the cohort taking GCSE mathematics in 2013 where more

than a fifth of pupils were certificating early (Ofqual, 2013c). Early entry is often associated

with lower grades (Ofsted, 2013) and an increase in the number of re-sits needed to achieve a

grade C (Noyes, Drake, Wake & Murphy, 2010). However, if a student achieved a grade C

early, they could, for example, undertake a qualification in a different subject (such as

statistics, having certificated in mathematics), which increased the overall points score per

student which affected some league tables. Schools could also redirect curriculum time for students to focus on subjects where they were not working at a C grade level yet. Both these activities might be seen as 'tactical behaviours' or 'gaming'. Finally, some schools also entered pupils for more than one qualification in the same subject. For example, 15% of pupils who certificated in mathematics in 2012 also entered at least one unit in a different qualification in the same subject (Ofqual, 2013d). Most candidates entering more than one qualification in mathematics were on the C/D borderline. As well as GCSEs, schools could choose to offer level 1/2 certificates (commonly known as International GCSEs), BTECs or other qualifications considered equivalent which would also count in the performance tables. This evidence was used by the then government to rule that, from September 2013, only the first entry or certification from a pupil in a subject would count towards a school's performance measure.

Another feature of GCSE qualifications was the use of tiered examination papers, which gave access to different ranges of grades at foundation (grades C to G) and higher (grades A* to C) level. Gilborn and Youdell (2000) suggested tiering entry practices are sometimes strategic, and Strand (2012) found evidence of systematic bias in the ways that students were entered for tiers in mathematics and science GCSEs. There is a long-standing belief among teachers that it could be easier to gain a grade C in mathematics at the higher tier (because of the number of marks needed for a grade C on the different papers), which also contributed to decisions about tier entrance (Taylor, 2016). In the legacy mathematics qualifications, tier entry choices for assessments was based on prior attainment and expected achievement (Dunne, Humphreys, & Sebba, 2007) and a majority of candidates were entered for the higher tier (around two thirds; Ofqual, 2017). The reformed qualifications also included changes to the availability of different tiers of examination. Tiers were removed in most subjects, but mathematics (as well as science and languages) continued to be assessed

7

through a foundation and higher tier, although the alignment or overlap between the grades available in the two tiers is different.

Coursework was introduced in 1988 within the newly created GCSEs. The aim was to assess skills that could not be assessed with a written exam. A task would be set by the teacher or the awarding organisation and candidates would carry out the task at home. Teachers would mark the task and the awarding organisation would carry out a moderation exercise on a sample of marked coursework to ensure consistency of marking across schools. In 2009, the decision was made to replace coursework with controlled assessments given concerns around the reliability and authenticity of coursework (Ofqual, 2013e:3). Controlled assessments are similar to coursework but they are carried out under controlled conditions. Finally, in 2015, the decision was made to reduce the amount of controlled assessment to a minimum. This decision was partly because pressures on schools to deliver good results led "to the preparation of controlled assessment to a point where the final work was not representative of a student's true level of replicable achievement, and sometimes also to over-marking, which in turn leads to unfairness to other students." (Ofqual, 2013f;7). In the reformed GCSE qualifications, most no longer include controlled assessments meaning that all assessments are externally marked. In August 2013, Ofqual also announced that the speaking and listening unit in English would not be counted as part of the qualification but would be graded and reported independently (Ofqual, 2013a), citing concerns over the 'fairness' of the results of the assessment and the consistency of the rigour of marking. Both sets of changes spoke to established fears about the reliability of teacher assessment, such as the influence of factors other than the quality of work on teacher judgements (Johnson, 2013).

For many years, schools had a great deal of choice about the type of qualifications for which they could enter their students, which included choices between modular or linear examinations. In 2010, the DfE announced that GCSEs should not be modular. After an

Ofqual consultation, it was announced that for certifications from June 2014 modular GCSEs would not be available in England and all assessments would be linear, meaning that all assessment for GCSE qualifications would be designed to be taken at the end of the course. Whilst this was mainly aimed at the assessments included in the new reformed qualifications that were to be accredited, it also applied to current qualifications, designed to be modular, which would be assessed at the end of the course until the new reformed qualifications were first taught in 2015 for English and mathematics (other subjects to be taught in the subsequent years). This also signalled the end of students being able to re-sit individual modules; the only way to re-sit a qualification is to re-sit all the assessments. Re-sitting is popularly linked with grade inflation but research evidence casts doubt on that relationship (Baird *et al*., 2013); Vidal Rodeiro and Nádas (2010) show that re-sitting more than one module is strongly associated with a lower grade outcome over all.

The key reforms which the teachers in this study referred to in the data presented below are, therefore:

- the disincentivising of early entry by making the first examination entry for any student 'count' for league tables
- the removal of tiers in English and the change in the balance of tiers in mathematics
- the removal of most controlled assessment.
- the change from modular to linear and the subsequent reduction in re-sit opportunities

In this paper, we present new empirical data which not only reiterates a tension between accountability and student need, but also concludes that the practices described as being for the good of the student were those very practices attributed to gaming when used by other schools, and that there is a fine line between using the system and playing the system.

Methods

The data used in this paper are a subset of data that were collected as part of a two-year

project entitled *Examination Reform: Impact of Linear and Modular Examinations at GCSE*.

The project is collaboration between the Oxford University Centre for Educational

Assessment (OUCEA) and Ofqual, investigating the effects of modular and linear

examinations upon outcomes and teacher practices at GCSE, and has taken a mixed-methods

approach. Data collection for the qualitative understanding of the perceptions of teachers in

relation to modular and linear examinations was undertaken in two phases. The first phase

was conducted in 2015, before the new linear qualifications came into effect in September of

that year (with the exception of three schools where interviews were conducted shortly after

this point), and the second phase was conducted in 2017, just before or as students sat the

new GCSE examinations for the first time (with the exception of one college where

interviews were conducted shortly after this point).

Sampling

In the first phase of research, the participating institutions in England were selected using a

random stratified sample divided equally by 1) institution type and 2) the kind of GCSE

examination entry policy they employed in mathematics in 2013 (linear or modular

qualifications, where available). If the institutions entered 50% or more of their GCSE

Mathematics pupils for a modular qualification in 2013, they were classed as having a

modular entry policy. Those that did not meet this requirement were classed as having a

linear entry policy. Care was also taken to ensure that the sampling only included those

schools that entered 50 or more candidates for GCSE Mathematics, with the exception of

special schools where those with the majority of their students entered for GCSEs were

identified. This led to a sample of 15 institutions, at which interviews were conducted

between April and November 2015, comprising three Further Education (FE) Colleges, six

academies, three secondary schools, two independent schools, and one grammar school. Of

the recruited institutions, seven had already been operating under the above-defined linear

policy.

In the second phase of the project, the same sample set of 15 institutions from Phase 1

was approached for a second round of interviews. Where institutions were unwilling or

unable to take part, appropriate replacement schools were approached for recruitment.

Twelve institutions were recruited (eleven of which were the same as in the first round),

which comprised two FE Colleges, four academies, two secondary schools, two independent

schools, one grammar school and one special school (we had been unable to recruit the

special school in time for the first phase of research).

Participants

During both phases, at each institution, a person with responsibility for English, a person with

responsibility for mathematics, and the Head Teacher were invited to participate in semi-

structured interviews. When alternatives had to be sought, the interviews were conducted

with suitable replacements, such as the Head of Key Stage 4 in each subject, or the Deputy

Head. On occasion, individuals were interviewed who served in dual capacity, and the

interview questions were directed to capture both responsibilities. Forty-three interviews

were conducted during the first wave and thirty-four interviews during the second wave,

giving a total of seventy-seven interviews.

Data Collection and Analysis

Semi-structured interviews were conducted during both phases of the research. In the first

phase, the interview schedule was designed drawing upon literature pertaining to the

changing GCSE examination structure. The interview schedule for the second phase of

11

interviews drew on the same literature but also on understanding of teachers' perceptions that was generated in the first wave of qualitative research. During both phases of the research, interviews were carried out by different members of the research team from both OUCEA and Ofqual, and the schedules served as a guide to discussion. During the first phase, an additional list of claims about modular and linear examinations generated from the literature was provided to the interviewers to prompt discussion when necessary.

All interviews from both phases were audio-recorded and then transcribed. Following transcription, in both phases, two researchers from the team coded six interviews using inductive and deductive techniques to generate a coding framework that was pertinent to the data in both phases. The coding frameworks were then used to code the whole datasets. Broad themes were established, and the transcripts were re-read for several instances of those themes. Each school was assigned a distinct numeric identifier, which will be used to reference them in the following analysis. Specific extracts from the data will reference the interviewee by job title and the school's assigned number, for example, *Deputy Director 7*.

Results

As will be seen from the results presented below, when talking about entry choices for students, the decisions are complex and this influences what is perceived by schools to be gaming or playing the system. In the first phase of research, decisions regarding modular entry, early entry, tiering and re-sits were all related to one another, meaning that there was not a single rationale for entry choices. During the second phase of research, findings about the changes in examination structure were confounded by curricular and grading changes. Since views about what constitutes gaming vary, we have focused on behaviours which were identified in the justification for the reforms, and on participants' own understanding of ethical and unethical behaviour in relation to GCSE examinations.

An evolving conversation: from playing the system to levelling the playing field

The practices and beliefs captured around entry choices in our interviews reflect the delicate balance that teachers strived to achieve in what they felt what was best for their students over what was needed for the best outcomes for the school.

> 'I think if schools can choose the best exam for their students, then they will try to do that. I think that is actually muddied by the accountability system, and I think that's where schools will put everybody into where they think the best results will come, not necessarily what will give the students the best outcomes.' (Deputy Head teacher 6, Phase 1)

The idea of 'playing the system' to improve the outcomes of the schools was introduced by teachers in both phases of our research, although the way that teachers discuss it has evolved over the two years. As the Head of Mathematics (School 6, Phase 1) noted:

> 'Personally, I think it's to stop schools playing the system. I think that it was felt that some schools would enter them for modular, then they'd re-sit, and they'd re-sit and re-sit, so that they could get the best grade possible, which of course is in the students' best interests, but I think perhaps people felt that schools weren't being accountable for not getting it right the first time'.

In the first phase of research this notion of 'playing the system' was explicitly talked about by seventeen interviewees; in the second phase of research, the discussion was more nuanced, and framed by fourteen teachers through the notions of fairness and a 'levelling of the playing field' for all institutions.

> 'I've always been very much, from the even playing field perspective, I am very much in favour of a linear course anyway, always have been. I think it's too easy to manipulate the system otherwise which, although we don't do, when you hear about everybody else doing it, it's just so unfair and you just can't manage it'. (Head of English 8, Phase 2)

13

A sense of the other

It is pertinent to note is that in both phases of this research, where teachers did speak about

playing the system, it was spoken of as something only **other** institutions did:

> '….I think there was quite a lot of playing around by some schools where, doing AQA
> modular with the same cohort, and then Edexcel linear to just really try and capitalise on
> that, and I think when you start getting to that level, you're losing sight of what we're
> here to do.' (Head of Mathematics 4, Phase 1)

The teachers in the first phase of interviews spoke of how they 'weren't a school who did

multiple, multiple, multiple re-sits' (Head teacher 1, Phase 1), or one of those schools who

had 'manipulated it [coursework] so much, that perhaps it's an unfair picture now' (Head of

English 4, Phase 1), or even one of those schools that on using early entry had 'high ability

students being allowed to just sit with a C …..[because] that's good enough [for their results]'

(Deputy Headteacher 6, Phase 1).One school however, did say that it felt like they had

'played the game a little bit' by entering different sets of their students for different

qualifications (Head of Mathematics 9, Phase 1) to maximise student outcomes; an indicator

of the tensions that teachers felt while trying to work towards the best outcomes for their

students.

> 'we really felt, although morally, doing first and then a second entry might be better for
> students, because schools in our position are so judged by the league tables & by
> performance tables, we just felt we couldn't do that.' (Deputy Head of Curriculum 2,
> Phase 1)

This sense of the 'other' was echoed in the second round of interviews as well, where

teachers spoke of having a level playing field in that modular qualifications had allowed other

institutions to not only 'cherry pick and retake and do what [they] want to do' (Vice

Principal, Mathematics 2, Phase 2), or 'teach to the test' (Head of English and Mathematics

14, Phase 2); but also using the flexibility to choose qualifications such as the 'the

international GCSE qualification was used at that C borderline, C borderline because it was

just seen as an easier [way to get a] C grade' (Head of English 15, Phase 2), or to choose

between modular and linear routes as to 'repeat modules and all that came with that…. to us

wasn't giving credit to those who could do it' (Deputy Headteacher 9, Phase 2). These

strategies were seen to create an unfair system that would disadvantage some schools over

others. 'There have been a few schools that abused what modular allowed you to do'

(Headteacher 5, Phase 2), and where 'other schools were using modular exams to their

advantage' (Deputy Head of Academics 8, Phase 2).

Given the practices that teachers have identified as 'gaming', it is important to look

closely at decisions governing teacher choices, and how they are justified in terms of doing

best by the students regarding choice of different examination boards, modular examinations,

early entry, and subsequent potential re-sits, and even for tier entry. Examination entry was

left to the individual departments for the most part, and so in certain schools the English and

mathematics examinations would have students entered in what made sense for the

departments and were not necessarily consistent across the school.

Changes in examination structure as a response to playing the system

Six interviewees explicitly felt that the changes in examination policy including the shift

exclusively to the linear examination structure was due to a perception on the part of policy

makers that the modular structure was 'making it too easy' (Head of Mathematics 5, Phase

1), or that the 'the league tables are not showing the results that the politicians think they

should be showing because some schools are playing the system' (Deputy Headteacher 9,

Phase 1). There was a suggestion that the changes to examination structure were put forward

by policymakers to 'stop schools playing a system as opposed to being a benefit of the

students' (Deputy Headteacher 6, Phase 1), and that teachers were just 'fiddling the grades and cheating and doing all these underhand things' (Head of English 2, Phase 1).

This sense of policy response to playing the system was echoed in the second phase of interviews through the way teachers felt that the new examination structures had helped restore a sense of fairness in the system. Five teachers shared the notion that having linear examinations was fairer, as 'everyone was [now] in the same boat' (Vice Principal of Curriculum 2, Phase2). This was made particularly obvious when teachers highlighted challenges with the modular approach, such as the 'unevenness' of controlled assessments at different schools, and the fact that there was the potential for abuse of the availability of retakes:

> 'No, I mean I suppose you...I mean one thing I would say it's probably going to bring a fairness back to the system, because you don't …you know if people are doing coursework in other schools and you start … you just don't know.' (Headteacher 4, Phase 2)

Choosing different qualifications for entry and a sense of autonomy

In the first phase of research, ten interviewees spoke of autonomy and a high degree of flexibility to choose a qualification that was the best for their students. In one school for example, all students were entered for a qualification that had coursework because having coursework was considered 'supportive for [their] students', with students' entry into different tiers based on both predicted performance and target grades (Head of English 2, Phase 1). At another, teachers could pick qualifications that suited the profile of their students, 'not necessarily in terms of result but in terms of engagement with curriculum' (Head of English 3, Phase 1), where one set of students was entered in for International GCSE, another for a GCSE responding to the needs of the different skill levels. Six teachers also spoke of how they used different qualifications for different attainment groups to help

provide their weaker students with the opportunities they would need to succeed;

> 'So yes, I feel the use of different courses for different groups at this school wasn't
> cynical; it wasn't chasing league table position, it was getting the best course for the
> students' (Head teacher 1, Phase 1)

Some teachers in this phase however, keenly felt that even though they approached qualification choice with the students' best interests in mind, it was often perceived by other schools and the general public that that they were chasing league table positions. As one teacher described it, their school held a 'moral value about the way [they'd] do things' (Head of English 3, Phase 1), and so has never felt the need to bring in different qualifications to play 'different cards on a school league table'. Another school highlighted the extent of the challenges teachers felt in their choice of qualification as follows:

> Well it was a number of things. I think it's part of the sort of pressure to sort of … you
> know to get better results really. And I mean at the time schools doing 'IGCSE' were
> getting much better English results, so we had to look into that. And I think the other
> thing that we found was that the speaking and listening still contributed to the overall
> mark, which they had just detached that, you see … so we thought the speaking and
> listening … I mean it suited our students (Head of English 7, Phase 1)

This sense of autonomy from Phase One was missed by some teachers in the second phase of research as reflected in five different schools. Teachers at these different schools felt that the choice of qualifications that could be used to 'tailor the approach for individual students' (Head of Mathematics 4, Phase 2) was now restricted to only a linear approach, providing very little real help for students. Five teachers spoke of having to balance the choice of qualification offered with the needs of progress measures, in that they were 'particularly [because of] the Progress 8 and Attainment 8 measures [having to] focus in more on certain subjects' (Deputy Head of Academics 8, Phase 2). This meant that there were fewer courses to offer students for their needs where 'there's much more resources now in the

core and in the EBacc subjects and you know things like Design Technology very rapidly are

sort of disappearing' (Headteacher 3, Phase 2).

'They have to do GCSEs, so the way that it's measured means that they have to sit the
GCSE. We can offer them a different qualification alongside it, but they have to sit the
GCSE and they have to sit both Language and Literature, um, and that is all driven by
performance measures, but the alternative is we drop in performance measures and then
the school ends up being looked at in far more detail, there's far more pressure.' (Head of
English 4, Phase 2)

It is interesting to note that in this phase, two institutions were still offering other

qualifications such as the International GCSE as the 'right specification hasn't come up for us

[in terms of suitability for their students]' (Head of English 14, Phase 2), even though it

meant their EBacc profiles were incomplete.

Choosing the modular approach

In both phases of our research, whilst the effect of entry into a modular or linear route was

not linked to membership of defined groups of students, some teachers did highlight why

choosing modular courses was of benefit to some students:

'You know we have to have exams that cater for all students, not the highest ability. You
know we can't have a system that causes a whole bunch of kids to fall by the wayside.'
(Head of English 4, Phase 2)

In Phase One, some teachers held the belief that the inability to enter students into a

modular course would disadvantage certain students, mainly 'from more vulnerable

groups…there are potentially going to be more students with no qualifications because of the

changes that have been made' (Deputy Headteacher 6, Phase 1). A number of teachers

mentioned worrying about the economically disadvantaged students, whom they felt needed

smaller chunks of examinations as they do 'struggle with resilience, struggle with retention of

information' (Deputy Headteacher 2, Phase 1). This concern was echoed in Phase Two of our interviews, the inability of students to take 'bite size exams' (Head of English and Mathematics 10, Phase 2) was criticized by thirteen teachers particularly when talking about the impact on students attracting Pupil Premium funding (disadvantaged students, often with low socioeconomic status). The 'all or nothing culture', created by the 'high stakes, final, terminal exams' (Headteacher 5, Phase 2) were cited as having a disproportionately negative impact on those Pupil Premium students who wouldn't have the best attendance (Head Teacher 3, Phase 2), supportive family structures (Head of English 7, Phase 2) or the cultural capital (Head of English 15, Phase 2) that would be needed to deal with the extremely stressful situations created by the terminal examinations.

> 'I think in terms of the setup, a modular route was better for the lower and middle ability students.' (Deputy Headteacher 14, Phase 2)

Modules and therefore modular exams were considered better for lower attaining students (Head of Mathematics 9, Phase 2), and there was a growing concern that these students would find it very challenging to cope when doing 'pure linear in two years' time' (Deputy Head Academic 8, Phase 2). In the first round of interviews, numerous interviewees stated that modular examinations had, in fact, been used as a way to motivate and engage students as there was never an 'examination more than two or three months away' (Head of Mathematics 7, Phase 1). This was reiterated in our second round of interviews where some teachers expressed concern with the rising disengagement and lack of motivation that students showed due to the linear structure, which was confounded by the 'hardness' of the new examinations (Vice Principal of Curriculum 2, Phase 2).

Six teachers in our second round of interviews also pointed out that having the ability to enter those students with special educational needs into modular examinations, provided

them with the ability to take the pressure off them, and manage their expectations (Head of Mathematics 15, Phase 2)

> 'Yeah. But I think overall, I think of the two systems, a system that had different ways of
> assessing seems fairer than a system that is completely linear and completely based on
> the same method of assessment because if your strength is not that, for example, if you're
> dyslexic then you can't succeed'. (Head of English 2, Phase 2)

Teachers described using early entry to modules to motivate their students; where early success was viewed as a highly motivating factor (Deputy Headteacher 2, Phase 1), or an early module failure could provide a 'reality check' (Deputy Headteacher 9, Phase 1) to a student who needed the encouragement to work harder. One interviewee described having 'two shots at the exam' to 'build up [their students'] confidence and build up their resilience' (Deputy Head of Curriculum 2, Phase 1). At one FE college, entry 'to modules was manipulated [based] on the strength of when a student was ready to enter the examination' (Headteacher 12, Phase 1).

Early entry and re-sits

It is pertinent to note that the discussion around decisions for early entry and re-sits was evident largely in the data from our first phase of research. The new qualifications were first offered for sitting in 2017, so early entry had not been possible for the current Year 11, and the changes to league-table rules (in which only the first examination result counts) have provided a clear policy steer for schools away from early entry. Teachers spoke of early entry in two ways – one where students were entered for certain modules early (as discussed in the section before) and the second where students sat the entire qualification early (at the end of Year 10).

> 'We had a policy of early entry in November for English and maths, which we held onto
> last year despite the pressure to give it up. I believe we used that very positively, and in

both our last two Ofsted reports, it's been recognised that we used it very positively.'

(Headteacher 1, Phase 1)

Teachers used early entry where they felt it was important for students to get 'one thing out of the way' (Head of English 6, Phase1) to alleviate the pressure students felt at the end of year 11, or where they worried for the 'futures' of vulnerable students (Deputy Head 7, Phase 1). On the other hand, there were three teachers who believed that allowing their students to sit early was risky, as they might not have been ready or even mature enough to take their examinations. The link between maturity and early entry was reiterated in our second phase of interviews where seven teachers felt that the move to linear was better for their students, as they genuinely felt that 'students, the older they get, the better they get…. Even [if they were] doing sort of early entry things and things like that, [they were] not benefitting the student at that point' (Vice Principal in charge of Mathematics 2, Phase 2).

Interviewees at five different institutions reported that change in government policy caused them to re-visit their early entry policies, which in turn had a knock-on effect on their re-sit policies. In one school, school leaders had taken the collective decision to continue offering their students early entry as that early entry and subsequent re-sit was the 'best chance for the kids getting a C grade' (Head of Mathematics 3, Phase 1). They were willing to see a slump in their official numbers to ensure that they were doing the best for their students. At another school however, where the school had historically offered an early entry policy in the 'best interests of students and their learning needs', the early entry option was no longer offered as 'schools in [their] position [were] so judged by the league tables and by performance tables, [they] felt that [they] just couldn't do that anymore' (Deputy Head of Curriculum 2, Phase 1).

Teachers described using their re-sit policies to provide a means for their students to improve their grades, but also as a tool to ensure they provided the chance for students to at

least achieve their desired C grades (Head of Literacy 1, Phase 1), something that was described as a means of 'gaming' in other conversations. Three teachers talked specifically of using their re-sit policy to ensure their students' C grades.

'Looking at the early entry and the schools that just put lots of...as many of their students through as possible, and high ability students being allowed to just sit with a C, to then say, 'Right, that's done; let's move on to the next, and you don't do English in year 11, or you don't do maths in year 11 because you've just got a C; that's good enough for us'. (Deputy Head 6, Phase 1)

For some teachers, re-sits provided a means of motivation and support to those who needed more intervention, where for others it provided a second chance to get it right 'like a driving test' (Head of Mathematics 15, Phase 1). From wanting to get the 'best results [they could] for their students' (Head of Literacy 1, Phase 1), to providing a way to take the 'pressure off them' (Head of Mathematics 2, Phase 1), four teachers described a generous policy of re-sits where any students who wanted to could re-sit their examinations, whilst the majority of institutions described a more discerning policy where teachers, and or school leaders, would sit down with students and really determine the need for a re-sit. Some teachers also talked about not having a re-sit policy at all in order to prevent their learners from saying 'I'll come back next year and do it' and not put in their maximum effort (Head of Mathematics 11, Phase 1), or having a re-sit policy for only those students who came in to Year 12 needing to re-take their examinations to obtain their required C grades (Head of English 4, Phase 1).

Choosing the right tier

During the first phase of research tiering was available in both English and mathematics. Both English and mathematics teachers described a complex set of decisions in determining what they perceived to be best for the student versus what needed to be done to ensure that a

student obtained at least a C grade. The practices described by teachers in entering students for a higher tier as they would have to do less to get a C grade or even entering all students for a foundation tier 'to consolidate a C' (Head of English 3, Phase 1), could arguably fit in their descriptions of other schools 'gaming' the accountability system. Six teachers in English and three teachers in Mathematics spoke of such practices;

> Respondent: So we mainly put in for higher. We only have maybe 19 or 20 or so in foundation, purely because they seem to be able to get a higher grade in the higher tier paper, even though they're getting fewer marks.
> Interviewer: So, you put a lower ability into the higher tier because the marks they need will be less and maybe better?
> Respondent: Yes. I mean, yes, it's the end result for them, even though maybe the exam is not so much fun. (Head of Mathematics 2, Phase 1)

> '…you know governmental policy on only taking their first entry. To me you know that has to be ignored for the good and for the sake of the students.' (Deputy Head 7, Phase 1)

Teachers also described using the different tiers to help students achieve their target grades; using the foundation tier to give their students confidence to then move them onto the higher tiers; and using a system of mock internal examinations throughout the year to ascertain which tier would be apt for which tier irrespective of which set they may have started out in.

During the second phase of research, the removal of tiering in English was welcomed by some teachers in that they could teach the same content to all, and not to have to argue 'with people about whether or not a foundation or a higher would be better' (Head of English 7, Phase 2); they were also cognizant that this removal of tiering was impacted by the change in curricular content. Six teachers felt that whilst it was a worthy intention noteworthy to 'teach[ing] to the top', there needed to be a 'reinforcing for the bottom' (Head of English 9, Phase 2), which they felt would not be possible with the new curricular demands.

Discussion

The examination system immediately before the first phase of this research included a great deal of flexibility in the kinds of exams schools could choose to enter their students in. The examination boards offered both linear and modular examination structures, students could be entered for both a modular and a linear exam in the same subject (either by means of being withdrawn from the modular before the final examination and entered for the same qualification taken linearly, or by means of entering parallel qualifications offered by different awarding bodies, such as international GCSE and GCSE), students could take qualifications before the age of 16 and these were included in the accountability measures, and students could and did retake exams multiple times. Many of these choices are no longer possible in the current examination system or the results do not contribute to performance table outcomes. So, whilst students can sit qualifications early and re-sit qualifications (but only whole qualifications, not individual modules), only the results of the first certification feature in the schools' results. Modular examinations are no longer an option at all. The 'game' that teachers and schools are playing has thus significantly changed.

The choices that schools and teachers make were justified by the teachers in this study in terms of being what they thought was best for their students. These justifications are not consistent in their use between schools or even between the old and the new examination systems. For example, some teachers spoke of the motivating effect of modular exams in giving students a reality check to encourage them to work harder, whilst others talked about the demotivating effect of repeatedly getting low grades in these exams. Playing the system was something that other schools did, yet many of the practices described as playing the system were spoken of in terms of the benefits to both the school and the students by those schools that used them.

All the teachers spoke about doing what is best for their students, but policy decisions varied considerably between schools and even between subjects within the same school. Decisions around examination entry not only considered student outcomes in terms of the grades they would get, but also the effects on motivation. The balancing is not just between school and student outcomes, but also between different types of outcome, not all of which affect performance tables.

With the removal of many of the examinations entry choices that schools can make, the system is now perceived to be more of a level playing field. However, only the options of modular examinations and coursework have been removed entirely from English language and mathematics qualifications. Choices around early entry and re-sits are still present but are disincentivised by the accountability measures used in England. Yet most schools reported that they are no longer using early entry or multiple re-sits with their students. Decisions that were justified as being best for students before the examination changes are no longer being made, in the context of rules where those decisions would disadvantage the school.

The practices described by the teachers in this study, both those the schools themselves used and those attributed to schools gaming the system, were within the rules of the system at the time. There appears to be fine lines between using the system, playing the system, manipulating the system, gaming the system, fiddling the system and cheating the system, often largely determined by the perspective of the individual as player or observer of the game.

Conclusion

In this paper, we have considered the ethical and educational tensions over examination entry policies and practices over a period of intense change. The 'game' is to produce the best

examination outcomes for individual pupils and for schools as a whole; the rules which could be exploited in the past included: early and multiple entry; modular versus linear entry, including the withdrawal of students from an almost complete but failing modular entry in order to enter for a linear examination at the end of the course; re-sitting individual modules repeatedly; and decisions over which tier to enter individual students for. Teachers were and are aware of the tensions between competing priorities over examinations, and acutely aware of the impact of choices surrounding examination entry on students.

Decisions made by schools on examination entry practices were consistently justified as being what was best for their students, and many of these decisions also benefited the school in terms of its portrayal in performance tables.   Yet many of these same practices are also described as playing the system or gaming by participants in schools which made different decisions. The findings suggest examination entry choices require serious thought, and that tensions arise between the differing needs of the individual student and the needs of the school; accusations of gaming are felt strongly by teachers, and the decisions and practices of teachers are justified by reference to the needs of the student. Yet when the rules of the game change so that the same decisions carry the same benefits to the individual student, but do not directly benefit the school, different decisions are made.

This paper has shown how teachers' perspectives are influenced by both the freedoms to make decisions about examination entries and the use of accountability measures. Further research is needed to explore how these different decisions are actually affecting students, both in terms of their qualifications but also in terms of the other factors that teachers take into account when making the decisions.

References

Baird, J.-A., Ahmed, A., Hopfenbeck, T., Brown, C., & Elliott, V. (2013). *Research evidence relating to proposals for reform of the GCSE: OUCEA Report*. Oxford University Centre for Educational Assessment.

Beck, U. (1992) *The risk society*. London: Sage Publications.

Behn, R.D. (2003). 'Why measure performance? Different purposes require different measures'. *Public Administration Review*, 63(5), 586–606.

de Wolf, I.F. & Janssens, F. J. G. (2007) Effects and side effects of inspections and accountability in education: an overview of empirical studies, *Oxford Review of Education*, 33:3, 379-396, DOI: 10.1080/03054980701366207

Dunne, M., Humphreys, S., & Sebba, J. (2007). *Effective teaching and learning for pupils in low attaining groups*. Nottingham, UK: Department for Children, Schools and Families. Retrieved from http://dera.ioe.ac.uk/6622/1/DCSF-RR011.pdf

Figlio, D.N. and Getzler, L.S., 2006. Accountability, ability and disability: Gaming the system?. In T. J. Gronberg & D. W. Jansen (ed.) *Improving School Accountability (Advances in Applied Microeconomics, Volume 14)* Emerald Group Publishing Limited, pp.35 - 49

Gillborn, D., & Youdell, D. (2000). *Rationing education: Policy, practice, reform and equity*. Buckingham: Open University Press.

Gove, M. (2013) *Ofqual policy steer letter: reforming Key Stage 4 qualifications.* Retrieved

from

https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/529404

/2013-02-07-letter-from-michael-gove-reform-of-ks4-qualifications.pdf

Johnson, S. (2013) On the reliability of high-stakes teacher assessment. *Research Papers in*

Education, 28(1), 91-105.

NAO (2003). *Making a difference: Performance of maintained secondary schools in*

England. (National Audit Office, London). Retrieved from

https://www.nao.org.uk/wp-content/uploads/2003/11/02031332.pdf

Noyes, A., Drake, P., Wake, G., & Murphy, R. (2010). *Evaluating mathematics pathways:*

Final report. London: DfE.

Office of Qualifications and Examinations Regulation. (2013a). *Our announcement on*

speaking and listening assessments. Retrieved from

https://ofqual.blog.gov.uk/2013/09/04/our-announcement-on-speaking-and-listening-

assessments/

Office of Qualifications and Examinations Regulation. (2013b). *Early entry in summer 2012*

and summer 2013 GCSEs and IGCSEs. Retrieved from

http://webarchive.nationalarchives.gov.uk/20141111122703/http://ofqual.gov.uk/files/

2013-08-22-annex-c-early-entry-in-summer-2012-and-summer-2013-gcses.pdf

Office of Qualifications and Examinations Regulation. (2013c). *Multiple entry in*

GCSE/IGCSE English/English language and mathematics in summer 2012. Retrieved

from

http://webarchive.nationalarchives.gov.uk/20141111122715/http://ofqual.gov.uk/files/

2013-08-22-annex-b-multiple-entry-in-gcse-igcse-english-english-language.pdf

Office of Qualifications and Examinations Regulation.(2013d).*Review of Controlled Assessment in GCSEs*. Retrieved from https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/377903/2013-06-11-review-of-controlled-assessment-in-GCSEs.pdf

Office of Qualifications and Examinations Regulation. (2013e). *GCSE reform consultation*. June 2013. Coventry: Ofqual.

Office of Qualifications and Examinations Regulation. (2014). *Summer GCSE and IGCSE entries for England: Provisional figures April 2014*. Coventry: Ofqual

Office of Qualifications and Examinations Regulation. (2017). *GCSE mathematics: understanding schools' approaches to tiering*. Retrieved from https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/591000/2017-02-10_Report_on_maths_tiering.pdf

Ofsted. (2013). *Schools' use of early entry to GCSE examinations: Its usage and impact*. Manchester: Ofsted.

Riddell, S., & Weedon, E. (2006). What counts as a reasonable adjustment? Dyslexic students and the concept of fair assessment. *International Studies in Sociology of Education*, *16*(1), 57-73.

Smith, P. (1993) Outcome-related performance indicators and organizational control in the public sector, *British Journal of Management*, 4, 135–151.

Strand, S. (2012). The White-Black Caribbean achievement gap: Tests, tiers and teacher expectations. *British Educational Research Journal*, 38, 75–101

Sturman, L. (2003) Teaching to the test: science or intuition? *Educational Research*, 45, 261-273. DOI: 10.1080/0013188032000137256

Taylor, R. C. (2016). The effects of accountability measures in English secondary schools: early and multiple entry to GCSE mathematics assessments. *Oxford Review of Education*, doi:10.1080/03054985.2016.1197829

Vidal Rodeiro, C.L. & Nádas, R. (2010) *Effects of modularisation*. Cambridge, Cambridge Assessment.

West, A., Mattei, P. and Roberts, J. (2011). Accountability and sanctions in English schools. *British Journal of Educational Studies*, 59, 41–62.

Wilson, D., Croxson, B. and Atkinson, A., 2006. "what gets measured gets done" Headteachers' responses to the English Secondary School Performance Management System. *Policy Studies*, 27(2), 153-171.