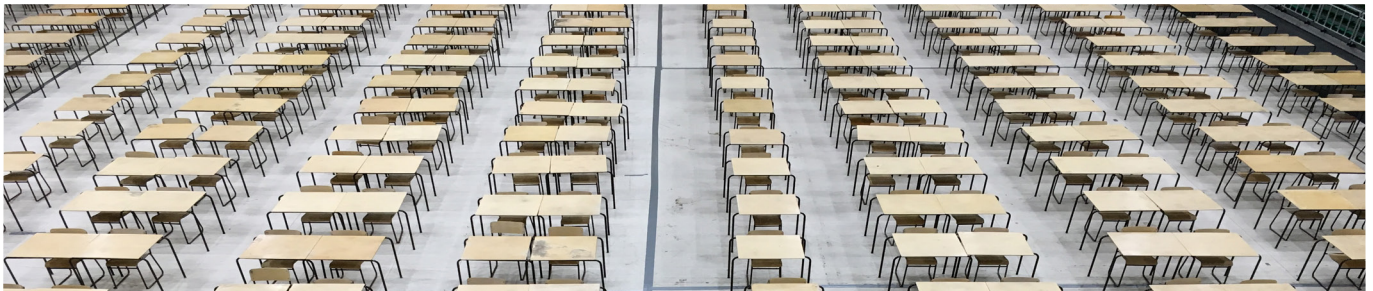




Briefing note: Should we abolish GCSEs?

Prepared by Gill Wyness



Summary

- Widespread disruption caused by Covid-19 resulted in the cancellation of GCSE exams for two years in a row in the UK. This has prompted some education leaders and policymakers to call for GCSEs to be abolished entirely.
- Those arguing for the abolition of GCSEs claim that high stakes testing results in a narrowing of the curriculum, and “teaching to the test”, and that exams put unnecessary stress on pupils.
- While academic studies agree that teaching to the test is a likely outcome of high-stakes testing, this may benefit students overall, especially those who may struggle with learning.
- There is little robust evidence of a causal relationship between exams and pupil wellbeing.
- Using coursework as an alternative means of assessment would also provide a ‘backup’ means of assessment if exams were cancelled again. But, coursework may measure different skills to exams and can favour certain pupils when it is internally set and graded.
- Abolishing assessment altogether could penalise students who leave full-time education at age 16.

Recommendations

- Policymakers should continue to assess pupils at age 16, to ensure all young people have a record of what they have learned across a broad range of subjects, before they take different pathways.
- Formal exams should continue to be used at age 16. Externally set and marked exams remain the fairest method of measuring pupil attainment, and there is little robust evidence of strong negative effects on learning or wellbeing.
- Externally set and marked coursework should be used alongside exam-based assessment to a greater extent than currently.
- In any reform of the system, policymakers should carefully consider what the alternatives to exam-based assessment are.

The Issue

The Covid-19 pandemic and resulting disruption to schooling led to the Government cancelling GCSE and A level exams in 2020 and 2021, with 2022's exams also likely to be disrupted. The cancellation of GCSEs, in particular, has renewed debate around the value of high-stakes testing, and prompted some politicians, practitioners and commentators to call for their permanent abolition.

The main objections to GCSE exams typically centre around two issues: 1) that high-stakes testing of this nature leads to narrowing of the curriculum with teachers having to “teach to the test” rather than provide a holistic learning experience, and 2) that this approach creates excessive stress for teenagers. Former Prime Minister John Major recently called for the abolition of GCSEs for these reasons, stating that “I have come to dislike these examinations due to the degree of stress and strain they impose upon students. Without the examinations, it would surely be possible to offer pupils a wider syllabus, providing a more rounded education.”

Teaching to the test

The academic literature exploring the consequences of high stakes exams tends to focus on academic accountability, since the two often go hand in hand, with test results used to produce school league tables. While not the focus of this briefing note, this literature tends to show that the introduction of standardised tests in high-stakes accountability regimes can be positive (William, 2010; Figlio and Loeb, 2011; Rockoff and Turner, 2010), though certain pupil types can gain more from accountability than others (e.g. Hanushek and Raymond, 2010 show that Hispanic pupils benefit most, and Black pupils least).

Lazear (2006) provides a theoretical analysis of this issue, arguing that high-stakes testing will always result in teaching to the test, because the incentives to learn the material and to teach it are so strong. The question, therefore, becomes whether teaching to the test is actually bad for students. Lazear makes

use of a helpful analogy between testing and speeding to provide intuition for his theory. We can think of the problem of how to prevent drivers from speeding. The question is, should the location of police speeding patrols be announced or kept secret from the public? If there are only few police and their locations are kept secret, the chances of getting caught will be low, so drivers will speed wherever they go. On the other hand, if the speeding patrols' locations are made public, at least speeding will be deterred at those locations.

The analogy is, if we teach a wide curriculum, with very open testing (e.g. where the exams involve a random draw of questions from a large body of knowledge), learners will have weak motivation to learn at all. On the other hand, if the questions are known (or are drawn from a narrow curriculum), then at least that part of the curriculum will be learned. Lazear also shows that the extent to which this is true depends on the type of learners in the population. For students for whom learning is high-cost (i.e. low attainers, who may struggle to learn easily), if testing is open, their incentive to learn will be weak, and they may learn nothing at all. But, if they are schooled in the questions in advance, at least they will learn something. Meanwhile, for low cost learners (high attainers, who find learning easier and more enjoyable), spreading incentives thin (by having a wider range of questions) will provide incentives to learn and will increase the amount of learning. This implies that – to the extent that it results in ‘teaching to the test’ – high stakes testing would be expected to narrow the gap between the top and the bottom of the attainment distribution.

Of course, there are many reasons to learn other than purely in order to pass the test in question. But at the margin this might matter for those with the same intrinsic motivation.

Furthermore, a small number of empirical studies highlight an important downside of high-stakes testing, which is that they may provide a “noisy” signal of pupil ability. This literature shows that pupils do worse in exams on hot days or during days with air pollution, and that this can go on to impact their future education-

al performance and even earnings (Ebenstein et al., 2016; Park, 2020).

Exams and pupil wellbeing

A further often cited objection to national testing concerns the mental strain it puts on young people (Weale, 2018). A large body of research has examined the link between school testing and pupil mental health, with some looking specifically at GCSEs. However, the vast majority of these studies (e.g. Putwain, 2009; Roome and Soan, 2019; Raey and Williams, 1999) are qualitative in nature, studying the experiences of pupils using focus groups and interviews with very small numbers of students. These studies imply a stress burden is experienced by pupils around exams, but are unable to establish any causal relationship between the two.

Some larger studies exist, typically using large scale survey data. For example, Denscombe (2000) analyses the stress related to the GCSE experiences of 15–16 year-olds in England, using a questionnaire survey of 1,648 pupils, as well as focus groups among smaller numbers. The study finds that young people believed GCSEs introduced an additional source of stress into their lives through the expectations of teachers and parents, and through internalised pressures to succeed. However, studies such as this rely on students' retrospective opinions about exams. Moreover, the majority of studies in this area have no meaningful control group who have not had to sit exams at the same time. Further important gaps remain, for example, few (if any) studies have attempted to evaluate whether pupils' mental wellbeing improves once the exams are over or the extent to which poor wellbeing during exam years may simply be picking up age effects.

Alternatives to high-stakes testing

While the current regime of high-stakes testing may have its weaknesses, any reform needs to consider what the alternatives are.

For the past two years – due to the abandonment of formal exams because of COVID-19

– GCSEs have been replaced with teacher assessed grades. There is currently no direct evidence on differences between “normal” exam-based assessment and teacher assessment at GCSE specifically (that the author is aware of). However, work by Burgess and Greaves (2013) examined this question for statutory tests at age 11. Their study, which uses administrative data on these age 11 (Key Stage 2) assessments – exploiting the fact that, at the time of their study, pupils were assessed both by their own teachers and by an externally set and marked exam – showed ethnic disparities in how pupils were judged by their teachers, compared to the judgements from “blinded” exams. In particular, they show that while 12.4% of white pupils received a grade from their teacher that was below their final test score, for black Caribbean pupils this possibility was 4 percentage points (around 32%) more likely. This, therefore, highlights the importance of externally set and marked exams in terms of fairness.

We can also draw evidence on the accuracy of teacher predicted grades for A level attainment (Murphy and Wyness, 2020). Evidence for 2013-2015 cohorts shows a high degree of inaccuracy of predicted grades, and that, for pupils who eventually went on to achieve AAB or more at A level, predictions were less generous for pupils from lower SES backgrounds and state schools. This may not be down to teacher bias, but rather that certain pupils have more volatile attainment trajectories, and so their eventual performance is harder to predict. Anders et al. (2020) attempt to predict the A level outcomes for pupils in England and Wales based on detailed data on their Key Stage 2 and GCSE outcomes (i.e. the sort of evidence teachers themselves would use to make predictions), finding that the A level results of high attaining pupils from state schools are particularly hard to predict.

The switch to teacher assessment is, of course, a temporary one brought in during a crisis and not one that has been suggested as a permanent alternative to the GCSE exam. However, one means of assessment, which is often discussed as a suitable alternative, is assessment through coursework. Since the

reforms of the 2010s, the use of coursework is quite rare for GCSEs and is only used in a small number of subjects.

There is little large-scale research on the viability of entirely replacing exams with coursework on pupil attainment, though a study by Benton (2016) looks at the question of how reliable coursework is, by comparing the extent to which GCSE coursework scores (versus exam scores) can predict future A level exam scores. The results of this correlation analysis show that GCSE scores from coursework appear to be more strongly associated with future A level achievement than GCSE scores from written examinations – in other words, coursework is often at least as predictive as externally marked tests in forecasting future performance. Of course, this does not tell us whether the underlying factor that both exams and coursework are explaining is knowledge or some other factor (that could be biased by other characteristics).

A report for Ofqual by Pinot de Moira (2020) touches on this question by looking at the extent to which different types of GCSE assessment (exams versus coursework) might be associated with different attainment levels among different groups of students. The study uses administrative data on GCSE takers over several cohorts, exploiting changes to assessment structure over time (i.e. where exams were replaced with coursework, or the weight given to coursework increased, etc.), to understand whether certain types of students respond better to coursework, or to exams, and is able to adjust for prior educational performance (Key Stage 2). The results find little evidence that students from different socio-economic status (SES) backgrounds perform differently in subjects with more coursework. There were some differences in performance by gender; after accounting for prior academic attainment at Key Stage 2, male students did better when the assessment was by exam and females did better when internally set and marked coursework was included in assessment. There was little evidence of ethnic disparities, although students of Chinese ethnicity performed worse when coursework was part of the assessment.

There are other issues regarding coursework with which we might be concerned. For example, there is likely to be more scope for parental or teacher interference in coursework and a greater possibility that pupils can cheat. There is little empirical evidence on these phenomena, though the rise of so-called essay mills in the undergraduate study (McKie, 2020) suggests there is scope for this. A final issue is that coursework is similar in nature to teacher assessment – so without external setting and marking it could suffer from similar biases and issues as those discussed above. Therefore, any move to coursework should involve external setting and external marking of coursework that counts towards a final grade. An additional advantage here is that in the event of future crises preventing exams, externally set and marked coursework scores would be readily available to fall back on.

Abolishing the GCSE would mean employers would have no standard way of comparing applicants who went on different pathways, while individuals who wish to change pathways later in life (e.g. who might wish to do A levels, having done a technical FE course) would have no standardised performance measure to demonstrate to a future employer or college their academic achievements from when they were at school.

Finally, the other alternative to the current GCSE regime is to simply abolish exams at age 15/16 altogether. GCSEs were originally set up as a terminal exam, to coincide with the school leaving age. However in 2015, the education leaving age was extended to age 18, meaning GCSEs are no longer a terminal exam. Nevertheless, it is important to note that GCSEs are the last point at which pupils are (broadly) assessed in the same way. After this point, pupils then go on to follow a variety of pathways. Around 45% of pupils do follow the traditional path to A levels, but many instead go on to technical pathways (Hupkau et al., 2016). Abolishing the GCSE would mean employers would have no standard way

of comparing applicants who went on different pathways, while individuals who wish to change pathways later in life (e.g. who might wish to do A levels, having done a technical FE course) would have no standardised performance measure to demonstrate to a future employer or college their academic achievements from when they were at school. Indeed, the abolition of GCSEs would mean that no one would have a marker of a core set of academic skills given that, post-16, everyone specialises in some way.

Summary

The Covid-19 pandemic, and resulting disruption to schooling, has meant that formal GCSE exams have not taken place for the past two years. This disruption has renewed calls by some education leaders and policy makers (including ex-Prime Minister John Major) to call for an end to GCSE exams altogether.

The typical objections to having high-stakes exams at age 15/16 are that they take away from a holistic learning experience, forcing schools to teach a narrow curriculum and incentivising teachers to “teach to the test”, and that they also negatively affect pupil wellbeing.

While academic theory does imply that teaching to the test is a likely outcome of high-stakes testing, students may actually benefit from this approach, particularly low attainers. Meanwhile causal evidence on the link between exams and pupil wellbeing is weak.

Alternatives to GCSE exams must be considered in any discussion about regime change. One potential alternative to exam-based assessment is to assess purely on coursework. This would provide a suitable back-up for exams, should there be a future crisis, and would mitigate the issue of pupils having a “bad day” on exam day. However, evidence shows that girls perform better than boys where coursework is used, suggesting that it measures different skills to exams. Moreover, if the coursework is internally marked, the evidence on teacher assessment suggests that certain pupils (e.g. Black Caribbean pupils) may receive lower grades compared to externally

marked exams, suggesting that any move to coursework should involve external setting and marking.

Finally, abolishing assessment altogether could penalise students who decide to change pathways in later years.

References

- Anders, J., Dilnot, C., Macmillan, L. & Wyness, G., (2020), Grade Expectations: How well can we predict future grades based on past performance?, No 20-14, CEPEO Working Paper Series, Centre for Education Policy and Equalising Opportunities, UCL Institute of Education, <https://EconPapers.repec.org/RePEc:ucl:cepeow:20-14>
- Blanden, J., & Gregg, P. (2004). Family income and educational attainment: a review of approaches and evidence for Britain. *Oxford Review of Economic Policy*, 20(2), 245-263.
- Burgess, S. M., Propper, C., Slater, H., & Wilson, D. (2005). *Who wins and who loses from school accountability? The distribution of educational gain in English secondary schools*, Bristol.
- Burgess, S., & Greaves, E. (2013). Test scores, subjective assessment, and stereotyping of ethnic minorities. *Journal of Labor Economics*, 31(3), 535-576.
- Benton, T. (2016). Evidence for the Reliability of Coursework. Cambridge Assessment, *AEA-Europe Annual Conference, Limassol*, Cyprus, 3-5 November 2016.
- Berliner, D. (2011). Rational responses to high stakes testing: The case of curriculum narrowing and the harm that follows. *Cambridge Journal of Education*. 41(3): 287-302. 10.1080/0305764X.2011.607151.
- Cho, E. Y. N., & Chan, T. M. S. (2020). Children’s wellbeing in a high-stakes testing environment: The case of Hong Kong. *Children and Youth Services Review*, 109, 104694.
- Denscombe, M. (2000). Social conditions for stress: Young people’s experience of doing GCSEs. *British Educational Research Journal*, 26(3), 359-374.

- Ebenstein, A., Lavy, V., & Roth, S. (2016). The long-run economic consequences of high-stakes examinations: Evidence from transitory variation in pollution. *American Economic Journal: Applied Economics*, 8(4), 36-65.
- Hanushek, E. A., & Raymond, M. E. (2005). Does school accountability lead to improved student performance?. *Journal of Policy Analysis and Management: The Journal of the Association for Public Policy Analysis and Management*, 24(2), 297-327.
- Hupkau, C., McNally, S., Ruiz-Valenzuela, J., & Ventura, G. (2017). Post-compulsory education in England: choices and implications. *National Institute Economic Review*, 240, R42-R57.
- Figlio, D. and Loeb, S. (2011). School Accountability. pp 383-421 in Eric A. Hanushek, Stephen Machin, and Ludger Woessmann (eds) *Handbooks in Economics* Volume 3, The Netherlands: North-Holland.
- Lazear, E. P. (2006). Speeding, terrorism, and teaching to the test. *The Quarterly Journal of Economics*, 121(3), 1029-1061.
- McKie, A (2020). *Essay mills 'targeting students' as pandemic crisis shifts HE online*. Available at: <https://www.timeshighereducation.com/news/essay-mills-targeting-students-pandemic-crisis-shifts-he-online> (accessed 09/08/2021).
- Murphy, R., & Wyness, G. (2020). Minority Report: the impact of predicted grades on university admissions of disadvantaged groups. *Education Economics*, 28(4), 333-350.
- Park, R. J. (2020). Hot temperature and high stakes performance. *Journal of Human Resources*.
- Pinot de Moira, A. (2020). *The impact of coursework on attainment dependent on student characteristics*. Coventry: Ofqual.
- Putwain, D. W. (2009). Assessment and examination stress in Key Stage 4. *British Educational Research Journal*, 35(3), 391-411.
- Reay, D. and Wiliam, D. (1999). I'll be a nothing structure, agency and the construction of identity through assessment. *British Educational Research Journal*, 25(3): 343-354.
- Rockoff, J., & Turner, L. J. (2010). Short-run impacts of accountability on school quality. *American Economic Journal: Economic Policy*, 2(4), 119-47.
- Roome, T. & Soan, C. A. (2019). GCSE exam stress: student perceptions of the effects on wellbeing and performance, *Pastoral Care in Education*, 37:4, 297-315.
- Weale, S (2018) Stress and serious anxiety: how the new GCSE is affecting mental health. Available at: <https://www.theguardian.com/education/2018/may/17/stress-and-serious-anxiety-how-the-new-gcse-is-affecting-mental-health> (accessed 28 July 2021).



Prepared by: Gill Wyness

Contact for further information:
Centre for Education Policy and Equalising
Opportunities (CEPEO)

www.ucl.ac.uk/ioe/cepeo

email: cepeo@ucl.ac.uk

Date: August 2021