Research Matters / 37
A Cambridge University Press & Assessment publication

ISSN: 1755-6031

Journal homepage: https://www.cambridgeassessment.org.uk/our-research/all-published-resources/research-matters/

Exploring speededness in pre-reform GCSEs (2009 to 2016)

Emma Walland (Research Division)

To cite this article: Walland, E. (2024). Exploring speededness in pre-reform GCSEs (2009 to 2016). *Research Matters: A Cambridge University Press & Assessment publication, 37*, 57–73. https://doi.org/10.17863/CAM.106035

To link this article: https://www.cambridgeassessment.org.uk/Images/research-matters-37-exploring-speededness-in-pre-reform-gcses-2009-2016.pdf

Abstract:
GCSE examinations (taken by students aged 16 years in England) are not intended to be speeded (i.e. to be partly a test of how quickly students can answer questions). However, there has been little research exploring this. The aim of this research was to explore the speededness of past GCSE written examinations, using only the data from scored responses to items from a sample of 340 GCSE components. Speededness was calculated as the average (mean) percentage marks lost from the longest string of unanswered items at the end of each student's examination paper. The potential impact of student ability on examination completion patterns was taken into account. The data suggested that most GCSEs analysed were unlikely to have been speeded. This method of exploring the speededness of exams using only scored responses has potential (although there are limitations), and it can flag potentially problematic components for further investigation.

Exploring speededness in pre-reform GCSEs (2009 to 2016)

Emma Walland (Research Division)

Background literature and research aim

The speededness of an assessment refers to the extent to which the assessment's time allocation influences students' performance, or the extent to which the assessment occurs under time pressure (Schnipke & Scrams, 1997). Speededness could be considered a source of increased demand[1] in an examination, along with many other sources that have already been explored (Ahmed & Pollitt, 2000; Crisp et al., 2008; Fisher-Hoch et al., 1997; Pollitt et al., 2008). Fisher-Hoch and Hughes (1996) proposed that demand can be valid or invalid. Valid demand is intended by the setter and related to the constructs being assessed, whereas invalid demand is unintentional, and can arise for several reasons. Speededness can negatively affect students' experiences of taking an assessment and, therefore, insufficient time allocation could be considered a source of invalid demand in an assessment that is not intended to be speeded. Skilled setters use their experience to determine proper assessment length; however, this can be a challenging task. The number and nature of items may depend on the age of students, the time available for testing, the type of items used and the type of interpretation to be made (Directorate for Quality and Standards in Education, Malta, 2022).

One way to estimate intended speededness is through the number of marks per minute of the assessment. The higher the number of marks per minute, the more speeded the assessment is likely to be. Whether the students experienced an assessment as speeded can also be explored by analysing not-reached items, or items left blank at the end of students' examination papers. Other methods look at response data to determine a point in the assessment where student performance deteriorated, as an indication of potential speededness (Shao et al., 2016). However, an important caveat is the use of "ramping" in some examination designs, whereby easier items are put at the start of the paper and the more difficult ones occur later. This means that items left unanswered at the end (or where student performance deteriorated) could also be due to students finding them too demanding. In England, conventional wisdom and research evidence indicate that GCSE students answer items sequentially (Spalding, 2011b) and GCSE examination developers across a range of subjects make use of ramping across

1 I use the term "demand" to refer to how challenging students find their examinations.

the examination paper (Johnson et al., 2017; Johnson & Rushton, 2019; Spalding, 2011a).[2] There is also evidence of ramping within each main question, with ramping occurring in each item that is part of an overall question (Johnson & Rushton, 2019). This must be considered when examining the items left unanswered at the end of a student's examination paper. Thus, speededness of an examination would be more strongly evidenced by higher ability students (rather than lower ability) leaving items blank at the end of a paper. The relationship for lower ability students can be complex. Pohl et al. (2014) argued that certain assessments that are highly speeded (e.g., reading tests) might have higher levels of omitting at the end for higher ability students than lower ability ones due to their different test-taking strategies. This could be because higher ability students may have worked more carefully on getting the items right whereas lower ability students may have skipped through the assessment quickly due to not being able to answer many of the items. They argue that the same pattern may not appear in less speeded assessments.

Other factors that could influence omit rates at the end are student motivation (Matters & Burnett, 2003; Pohl et al., 2014) and guessing. The former should not be a major concern for high-stakes examinations such as GCSEs. Regarding the latter, methods to detect guessing have been developed for multiple-choice assessments, where students are likely to engage in rapid non-systematic guessing (Schnipke & Scrams, 1997; Yamamoto, 1995), but these are not suitable for less constrained items where students are less likely or less able to guess and detecting guessing would be much more complex (Jones, 2019; Pollitt et al., 2008). Lastly, students' test-taking behaviour may be influenced by their personal characteristics. Matters & Burnett (2003) found that test-irrelevant thinking and academic self-concept predicted whether students were likely to omit short-response test items.

While GCSE written examinations are not intended to be speeded, there has been little research exploring this (as noted by Spalding (2011a) and Wheadon (2011)). Further investigation of this is important for assessing the fairness and validity of assessments. The aim of this research was to investigate the speededness of past GCSE written examinations, using a method that only considered the scored responses to items and whether they were omitted.

Method

Data

I selected a sample of 340 GCSE written examination components[3] for analysis. These components were from Physics, Science, Chemistry, Biology and Mathematics qualifications offered by OCR. These components were anticipated to have large entries as well as large numbers of items. None of the components in the sample had optional items. The main reason for that was that such items would show as missing in the data and, therefore, would confound the analysis.

2 Note that ramping is not suitable for all types of examinations, for example, English Language GCE, which only has a few long answer items intended to be of equal demand.
3 GCSEs are made up of separate exams or non-exam assessments called "components".

The time period for analysis was from 2009 to 2016, prior to the GCSE reform (Ofqual, 2018).[4] I included both higher and foundation tier components,[5] and excluded components with fewer than 100 entries from the analysis.

I conducted the analyses using SAS Enterprise Guide (version 7.1). For each component the following information was available: number of items, maximum mark for each item, and time allowed. For each student the data consisted of the scored response on each item, or an indicator of "missing" if the student had not attempted the item.

Prior to analysis I computed indicators of potential speededness for each component. These consisted of the marks per minute and the average (mean) percentage marks lost from the longest string of unanswered items at the end of each student's examination paper, referred to as "average percentage lost (at the end)". For each component, I also calculated: 1) the average percentage lost (at the end) per quartile of student achievement; and 2) the median percentage marks lost. The final dataset included all the above component-level data and was used for the main analyses.

Data analysis

First, I analysed the dataset descriptively in terms of marks per minute and the average percentage lost (at the end), and how this differed for different achievement quartiles and tiers. This enabled me to identify potentially speeded components. I only considered a component to be potentially speeded if the average percentage lost (at the end) was high for higher achieving students in addition to, or instead of, lower achieving students. Lower achieving students taking a highly speeded paper may not have a high percentage of marks lost at the end of their papers. This is because, theoretically, they could progress more quickly through their papers and reach the end due to not being able to answer many of the items earlier on in the paper (Pohl et al., 2014).[6]

Results and discussion

Of the total 340 components I analysed, there were 78 Mathematics, 78 Physics, 72 Science, 68 Chemistry and 44 Biology components. 170 of them were foundation and 170 were higher tier. Years ranged from 2009 to 2016, with 2012 being the most represented year. Table 1 shows the descriptive statistics for the numeric variables, rounded to two decimal places unless otherwise stated. Table 1 shows that the number of students in each component ranged from 137 to 55 564, with a mean of 14 185 (rounded to the nearest whole number). Examination duration ranged from 40 minutes to two hours, with a mean of just over an hour. The number of items per component ranged from 19 to 66, with a mean of 34 items.

4 The reform led to a new grading scale being used, among other changes.

5 Tiering is used in some GCSEs (e.g., Science and Mathematics) to better allow for the wide range of abilities at this level. For the assessments in this analysis, foundation tier components were graded C to U, and higher tier components were graded A* to E.

6 Student motivation can also be a factor influencing test completion. However, in this context, I assumed that student motivation was generally high as GCSEs are high-stakes examinations for students.

Table 1: Descriptive statistics for numerical variables (n=340 components)

Variable	Median	Mean	Std Dev	Minimum	Maximum
Number of students	11 666.50	14 185.18	13 295.70	137.00	55 564.00
Exam duration (minutes)	60.00	65.47	21.81	40.00	120.00
Total number of items	30.00	34.41	11.08	19.00	66.00
Maximum raw mark	60.00	64.63	18.69	42.00	100.00
Marks per minute	1.00	1.00	0.06	0.83	1.11
Average (mean) % lost (at the end)	0.65	0.81	0.65	0.01	4.05
Median % lost (at the end)	0.00	0.01	0.18	0.00	3.33
Average (mean) % lost (at the end) (Q0)	1.95	2.29	1.73	0.04	8.08
Average (mean) % lost (at the end) (Q1)	0.37	0.58	0.59	0.00	3.90
Average (mean) % lost (at the end) (Q2)	0.18	0.27	0.33	0.00	3.49
Average (mean) % lost (at the end) (Q3)	0.05	0.10	0.22	0.00	3.34
Average (mean) % completed	92.97	90.80	8.92	0.06	99.93
Average (mean) % not completed	7.03	9.20	8.92	0.07	99.94
Median % lost (at the end) (Q0)	0.00	0.09	0.55	0.00	5.00
Median % lost (at the end) (Q1)	0.00	0.02	0.24	0.00	3.33
Median % lost (at the end) (Q2)	0.00	0.01	0.18	0.00	3.33
Median % lost (at the end) (Q3)	0.00	0.01	0.18	0.00	3.33

Note. Q stands for Quartile, with 0 representing the lowest achieving quartile and 3 the highest. Average % completed and not completed refers to the percentage of students who completed or failed to complete their examinations, respectively. Where the term "average" is used, it refers to the mean.

Marks per minute

Table 1 shows that the mean marks per minute across all components was approximately 1, ranging from 0.83 to 1.11. This does not appear to be problematic in terms of speededness. For example, online guidance given to students in the context of GCSE History suggests that 1 mark per minute is a good rough guide to work towards (OCR, 2024). However, this does also depend on the nature of the items, for example, how much reading the item requires and how long it takes to produce a response. The distribution of marks per minute for each subject group is illustrated in Figure 1. The figure shows that most of the marks per minute across subject groups were around 1. There was more of a range in the data for Mathematics than for the other subjects.

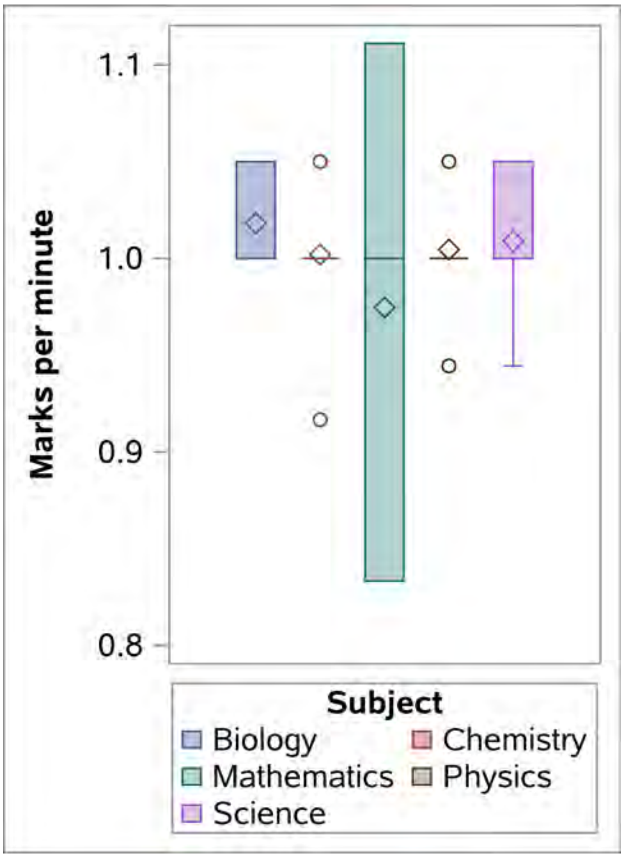


Figure 1: Schematic box and whisker plot showing the distribution of marks per minute for each subject group for all components. The diamonds represent the means, the box height represents the interquartile range, the horizontal line in each box represents the medians and the circles represent outliers. (The two subjects with no visible boxes had nearly the entire distribution clustered near 1 with only a few outliers).

There were 21 components with the highest number of marks per minute (1.1) and they were all Mathematics components. Marks per minute are about the **intended** speededness of the assessment, as it is determined at the design phase. The results of the analyses here indicate that the exams were likely not intended to be speeded based on the marks per minute. The average percentage lost (at the end), however, can indicate how the students **experienced** the assessment and whether it may have been experienced as speeded. I explore this subsequently.

Average percentage lost (at the end)

As shown in Table 1, the average percentage marks lost (at the end) by students across all achievement groups and tiers was 0.81 per cent (SD=0.65) and the median (of the medians across all achievement groups) was 0 with a range of 0 to 3.33 per cent. Across all components, an average of 90.80 per cent of students completed their examinations. This indicates that overall, there was little average percentage lost (at the end) across the sample. Figure 2 shows that the average percentage lost (at the end) was slightly higher for Mathematics components, and was very similar for the three single sciences (Biology, Chemistry and Physics).

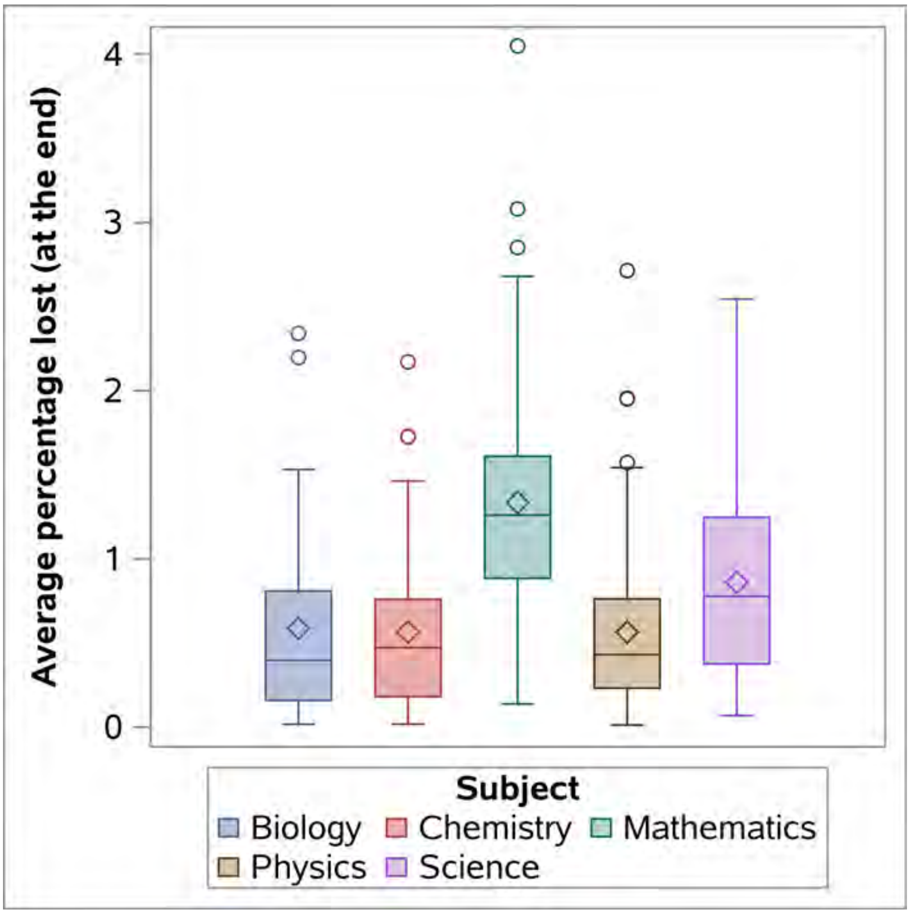


Figure 2: Schematic box and whisker plot showing the average percentage lost (at the end) for each subject group included in the analysis. The diamonds in the boxes represent the means, the horizontal lines within each box represent the medians, and the box height represents the interquartile range. The circles outside the boxes represent outliers.

I examined next the average percentage lost (at the end) for different student achievements and different tiers. The average percentage lost (at the end) for the highest achieving students in each component was much smaller, at 0.10 per cent. For the lowest achieving students it was much larger than for all students, at 2.29 per cent (Table 1). As noted, if a component was speeded, we would expect to see high levels of average percentage lost (at the end) for higher achieving students. Foundation tier components had higher levels of marks lost (at the end) in general compared with their higher tier counterparts, as shown in Figure 3. The difference between foundation and higher tier components decreased as student achievement increased (i.e., as we move from Quartile 0 to Quartile 3).

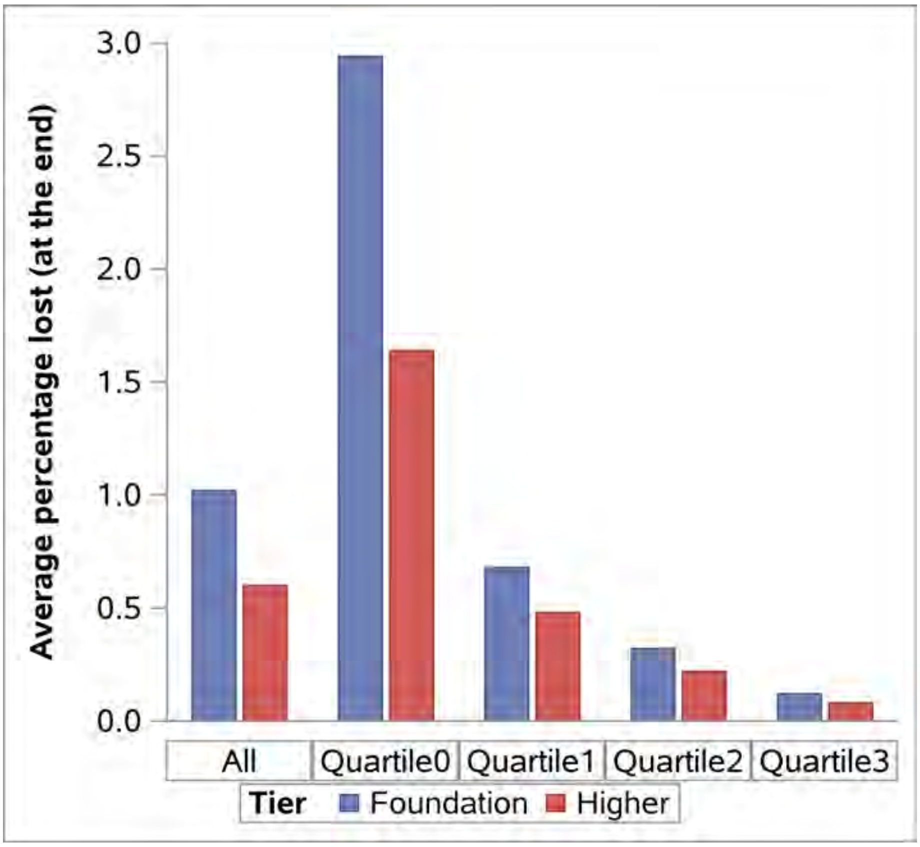


Figure 3: Average percentage lost (at the end) for each tier and each quartile of achievement, and overall

If an assessment was speeded, and students experienced time pressure, there is likely to be strings of items missing at the end for higher achieving students (Pohl et al., 2014). Given that there were few instances of omission at the end for higher achieving students, this suggests that most GCSEs were not speeded and items missing at the end were more likely due to student ability. Regarding low ability students, they likely experienced the items at the end as demanding and omitted them. But there are other possibilities: for example, that lower ability students are slower workers in general or have lower levels of motivation to complete their assessments. As noted previously, one theory according to Pohl et al. (2014), is that in a highly speeded assessment, lower ability students would omit fewer items at the end than higher ability students due to differences in test-taking strategies. I explored the data to investigate any instances where this pattern occurred, and no such examples were found.

The average percentage lost (at the end) for all components, for all students together (on the left) and for the highest achieving students (on the right), is illustrated in Figure 4. This shows that most of the components had a very low average percentage lost (at the end). However, there were some outliers with relatively high values, which could indicate speededness.

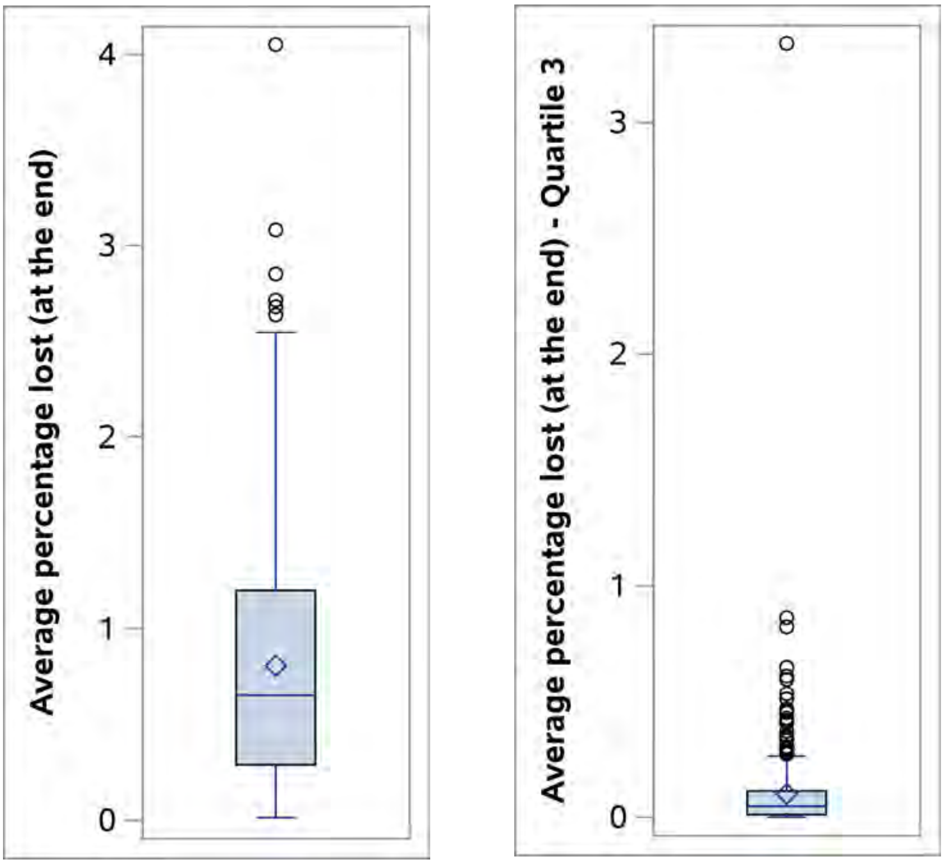


Figure 4: Schematic box and whisker plots showing the average percentage lost (at the end) for all components in the analysis. The figure on the left is the overall average percentage lost (at the end) across all students and the figure on the right is for the highest achieving quartile. The diamond represents the group mean, the box height represents the interquartile range, and the horizontal line within the box represents the median. The circles represent the outliers.

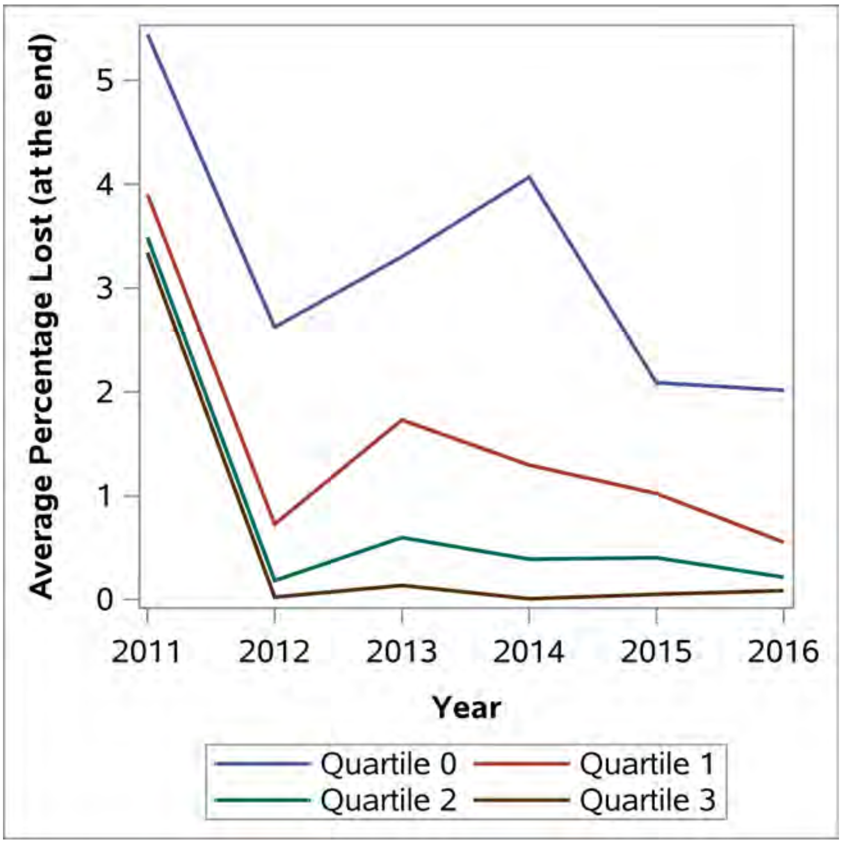
I looked at the components with the highest average percentage lost (at the end) for all students (Appendix A), as well as for higher achieving students (Appendix B). When all students were considered, there were 17 components which had an average percentage lost (at the end) above 2.00 per cent. Of these, the top six would be considered outliers according to Tukey's fences[7] (the upper bound was 2.57 per cent). For the students in the highest achieving quartile (Q3), 29 components were identified as outliers using Tukey's fences (in this case the upper bound was 0.26 per cent).

7 Tukey's fence is a method to detect outliers. Outliers are defined as values higher than Q^3 + 1.5(IQR) and lower than Q^1 – 1.5(IQR). Q^3 refers to Quartile 3, Q^1 to Quartile 1 and IQR to the inter-quartile range (Q^3-Q^1).

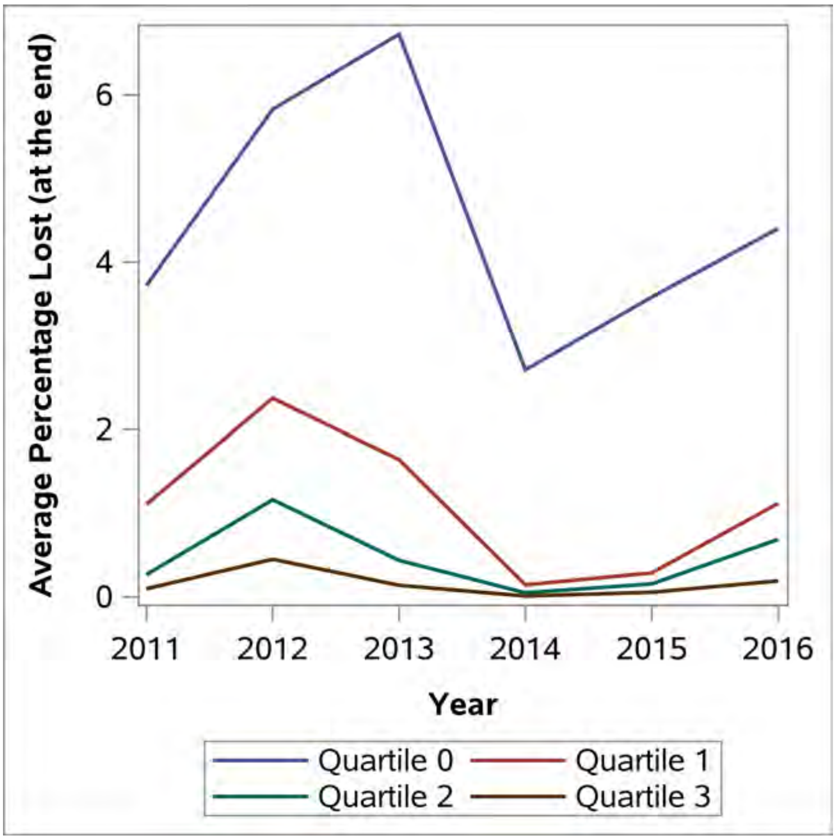
Analysis of the component with the highest average percentage lost (at the end)

The component with the highest average percentage lost (at the end) was the Mathematics M101 higher tier paper in 2011. The marks per minute was 1, which was the same as for the equivalent component in other years. This component also had the highest average percentage lost (at the end) for higher achieving students, at 3.34 per cent. In fact 99.94 per cent of the students who took this examination (11 527 students) did not complete it. However, nearly all these students did reach the penultimate item. The median last item attempted was the penultimate one and the median number of marks lost was 2, which was the tariff for the last item. As the exam was out of 60 marks, the last item was worth 3.33 per cent of the assessment. This suggests that students either ran out of time to reach the last item (slight speededness), or the last item was particularly challenging, even for higher achieving students, and so they omitted it. As the entry was large, the finding is unlikely to be related to the particular cohort.

I present the data for the Mathematics M101 component in the following graphs. Figure 5A shows the data for the higher tier and Figure 5B for foundation tier. Figure 5A shows that the average percentage lost (at the end) was higher in 2012 than in the other years, for all but the highest achieving quartile. Figure 5 shows that the percentage lost was highest for lower achieving students in all years and in both tiers. The number of students sitting this exam (across both tiers) ranged from 1439 to 15 148.



(A)



(B)

Figure 5: Average percentage lost (at the end) for higher tier (A) and foundation tier (B) students for Mathematics (M101)

Figure 6 shows the omit rate as a function of item position for Mathematics M101 higher tier in 2011. The omit rate for an item refers to the proportion of students who had no recorded response for that item. The figure shows that most items were completed by most students throughout the assessment, but that the final item was not completed by most students.

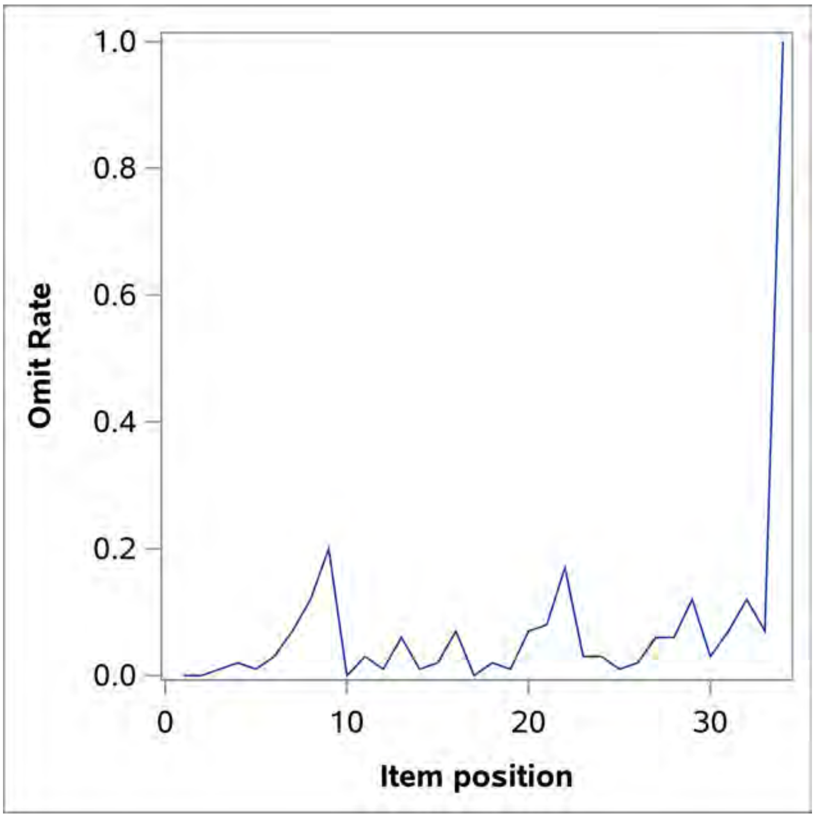


Figure 6: Omit rate as a function of item position, for Mathematics M101 higher tier in 2011

Figure 7 shows the facilities of each item, in the order they appear in the assessment. The facility of an item refers to the proportion of students who responded correctly to the item, which is generally used as an indicator of how easy an item is. The figure shows that the final item had a very low facility, which would usually indicate a demanding item (i.e., an item that a large proportion of students did not respond correctly to). However, as this statistic is based on the number of students who attempted the item but did not correctly answer it, it is not very informative in this case because only seven students attempted the item.

Figure 7 shows some evidence of ramping across the paper and within each item. The items later on in the test and the later items within each question were generally more difficult than the earlier ones.

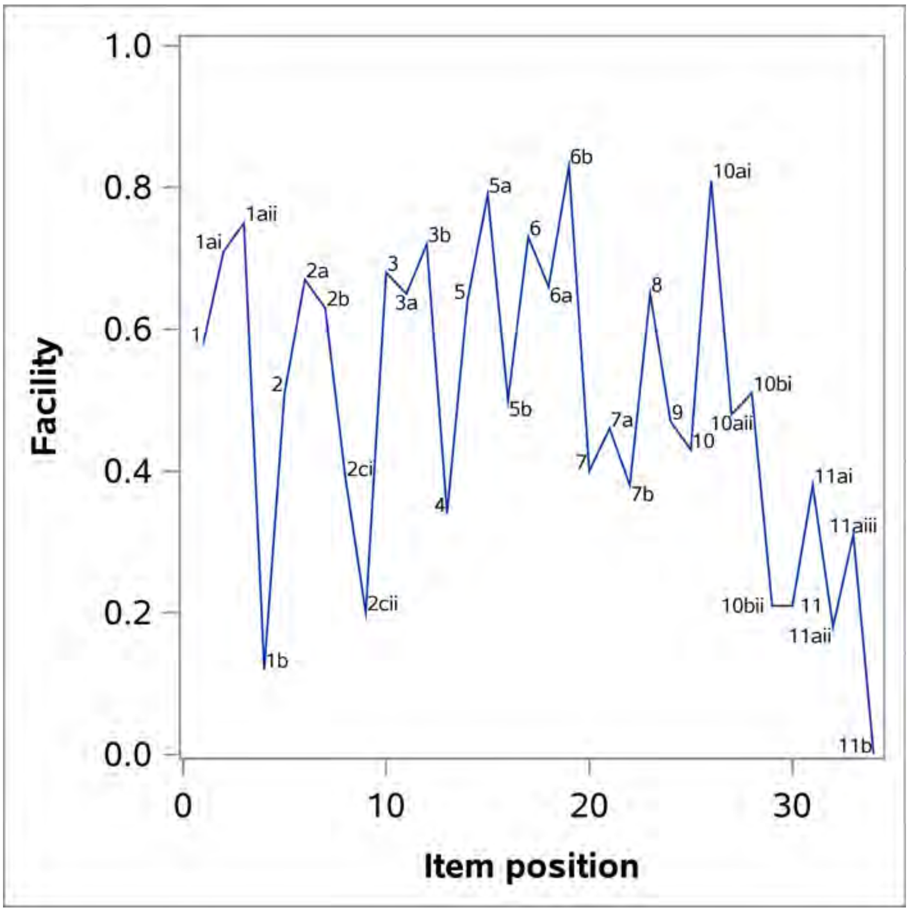


Figure 7: The relationship between item position and facility, for Mathematics M101 higher tier in 2011

Conclusion

The outcomes of this research show that there was little to no evidence of speededness in the pre-reform GCSE components analysed. Students did not seem to have been working under time pressure, according to the measure of speededness I used. This was indicated by very low levels of average percentage lost (at the end) for all, including higher achieving, students and also what appears to be appropriate marks per minute of around 1 mark per minute (OCR, 2024). This is reassuring as GCSEs are not intended to be speeded, and adds to our understanding of this under-researched area.

There were, however, some components, mainly from GCSE Mathematics, that showed some evidence of speededness. This could mean that students experienced time pressure, which could constitute an invalid source of demand, according to Fisher-Hoch and Hughes (1996). However, upon further investigation of the most outlying one it appeared that this was a result of most students not completing the last item. As there was only one item omitted, it is unclear whether students ran out of time or found it too demanding to attempt.

The limitations of using only data from scored responses (and omitted items in particular) to study potential speededness include that we are unable to account for student motivation and the role of test-taking strategies including guessing. The results are also complicated if there are higher tariff items at the end of an assessment, as students may start to write an answer but not complete it rather than omit it entirely. Speededness may also be indicated by students' performances deteriorating towards the end of a test, due to being rushed. Methods to detect this, such as change-point analysis (Shao et al., 2016), could be useful to complement the data on item omission. However, these methods may not function well in assessments with items that are ordered by increasing demand. The method I used is most useful in situations where motivation is high, items are not guessable, items are ordered by demand, students complete tests in order, there are no optional items, there are many items, and items each have small mark tariffs.

Bearing these limitations in mind, the research provides an example of a data-driven means of identifying assessments with potentially inappropriate time allowances using only data from scored responses. The method could be a useful tool to flag potentially problematic components which can then be investigated further. The data can be combined with other sources of data about speededness including post-administration surveys (see, for example, Steedle et al., 2022), and expert judgements about examination length. With the potential rise in computer-based tests, the data could also be used together with response time data in the future, to evaluate speededness.

Having methods to identify speededness can be useful for assessment designers and evaluators in relation to issues of validity and fairness relative to individual characteristics of students. Future research identifying whether particular students did not complete many of their examinations across different subjects and over time would also be interesting. This would lead to understanding whether running out of time is a stable personality trait that occurs across domains and over time, or whether it is specific to each assessment situation.

References

Ahmed, A., & Pollitt, A. (2000). *Observing context in action*. IAEA conference, Jerusalem.

Crisp, V., Sweiry, E., Ahmed, A., & Pollitt, A. (2008). Tales of the expected: The influence of students' expectations on question validity and implications for writing exam questions. *Educational Research, 50*(1), 95–115.

Directorate for Quality and Standards in Education, Malta. (2022). *Guidelines for paper setters: Educational assessment unit*.

Fisher-Hoch, H., & Hughes, S. (1996). *What makes mathematics exam questions difficult?* British Educational Research Association, University of Lancaster, England.

Fisher-Hoch, H., Hughes, S., & Bramley, T. (1997). *What makes GCSE examination questions difficult? Outcomes of manipulating difficulty of GCSE questions*. British Educational Research Association Annual Conference, York.

Johnson, M., Constantinou, F., & Crisp, V. (2017). How do question writers compose external examination questions? Question writing as a socio-cognitive process. *British Educational Research Journal, 43*(4), 700–719.

Johnson, M., & Rushton, N. (2019). A culture of question writing: Professional examination question writers' practices. *Educational Research, 61*(2), 197–213.

Jones, E. (2019). 'Multiple choice exams are easier than written exams!' *APMG International*

Matters, G., & Burnett, P. C. (2003). Psychological predictors of the propensity to omit short-response items on a high-stakes achievement test. *Educational and Psychological Measurement, 63*(2), 239–256.

OCR. (2024). *GCSE History A: How long should students spend on each question in the exams?*

Ofqual. (2018). *Get the facts: GCSE reform*.

Pohl, S., Gräfe, L., & Rose, N. (2014). Dealing with omitted and not-reached items in competence tests: Evaluating approaches accounting for missing responses in item response theory models. *Educational and Psychological Measurement, 74*(3), 423–452.

Pollitt, A., Ahmed, A., Baird, J., Tognolini, J., & Davidson, M. (2008). *Improving the quality of GCSE assessment.*

Schnipke, D. L., & Scrams, D. J. (1997). Modeling item response times with a two-state mixture model: A new method of measuring speededness. *Journal of Educational Measurement, 34*(3), 213-232.

Shao, C., Li, J., & Cheng, Y. (2016). Detection of test speededness using change-point analysis. *Psychometrika, 81*(4), 1118–1141.

Spalding, V. (2011a). *Is an exam paper greater than the sum of its parts? A literature review of question paper structure and presentation*. AQA.

Spalding, V. (2011b). *Structuring and formatting examination papers: Examiners' views of good practice*. AQA.

Steedle, J. T., Cho, Y. W., Wang, S., Arthur, A. M., & Li, D. (2022). Mode effects in college admissions testing and differential speededness as a possible explanation. *Educational Measurement: Issues and Practice, 41*(3), 14–25.

Wheadon, C. (2011). *An Item Response Theory approach to the maintenance of standards in public examinations in England* [Doctoral Thesis, Durham University].

Yamamoto, K. (1995). Estimating the effects of test length and test time on parameter estimation using the HYBRID model. *ETS Research Report Series, 1995(1)*, i-39.

Appendices

Appendix A – Components with the highest average percentage lost (at the end), all students

Code	Component name	Year	Average % lost	Average % lost Q0	Average % lost Q3	N students	Tier
M101	**Mathematics Unit B**	**2011**	**4.05**	**5.44**	**3.34**	**11 527**	**H**
M100	**Mathematics Unit A**	**2016**	**3.08**	**7.19**	**0.61**	**1635**	**F**
M103	**Mathematics Paper 1**	**2011**	**2.85**	**6.11**	**0.86**	**1194**	**F**
P100	**Physics A Modules P1, P2, P3**	**2012**	**2.71**	**7.20**	**0.21**	**3901**	**F**
M108	**Mathematics B**	**2013**	**2.68**	**7.48**	**0.11**	**17 103**	**H**
M100	**Mathematics Unit A**	**2011**	**2.64**	**5.94**	**0.51**	**6490**	**F**
S107	Science B: Unit 2 (B2, C2, P2)	2009	2.54	5.12	0.82	34 910	F
M101	Mathematics Unit B	2012	2.46	5.83	0.45	10 970	F
S109	Science Modules B2, C2, P2	2016	2.42	8.08	0.11	15 265	F
M108	Mathematics B	2013	2.40	6.46	0.10	17 155	H
S108	Science Modules B1, C1, P1	2012	2.39	5.92	0.65	14 203	F
B101	Biology A Modules B4, B5, B6	2012	2.34	7.84	0.37	1105	F
M108	Mathematics B	2014	2.29	5.22	0.42	30 484	F
B100	Biology A Modules B1, B2, B3	2012	2.20	5.99	0.46	3911	F
C105	Chemistry A: Unit 3	2010	2.17	6.81	0.15	634	F
M101	Mathematics Unit B	2013	2.17	6.72	0.14	1680	F
M107	Mathematics A	2009	2.06	5.28	0.17	22 963	F

Note: The six components that are outliers according to Tukey's upper fence are marked in bold. The assessment codes were created by the researcher and are not the real codes.

Appendix B – Components with the highest average percentage lost (at the end) for higher achieving students (Quartile 3)

Code	Component name	Year	Average % lost	Average % lost Q0	Average % lost Q3	N students	Tier
M101	Mathematics Unit B	2011	4.05	5.44	3.34	11 527	H
M103	Mathematics Paper 1	2011	2.85	6.11	0.86	1194	F
S107	Science B: Unit 2 (B2, C2, P2)	2009	2.54	5.12	0.82	34 910	F
S108	Science Modules B1, C1, P1	2012	2.39	5.92	0.65	14 203	F
M100	Mathematics Unit A	2016	3.08	7.19	0.61	1635	F
P100	Physics A Modules P1, P2, P3	2015	1.95	4.12	0.59	50 096	H
P107	Physics Modules P1, P2, P3	2012	1.46	3.58	0.53	843	F
M100	Mathematics Unit A	2011	2.64	5.94	0.51	6490	F
S105	Science A: Unit 3 (B3, C3, P3)	2012	1.90	4.49	0.47	6504	F
C100	Chemistry A Modules C1, C2, C3	2012	1.73	4.63	0.46	4194	F
B100	Biology A Modules B1, B2, B3	2012	2.20	5.99	0.46	3911	F
M101	Mathematics Unit B	2012	2.46	5.83	0.45	10 970	F
P104	Physics A: Unit 2 (P4, P5, P6)	2012	1.20	2.05	0.43	12 009	H
M108	Mathematics B	2014	2.29	5.22	0.42	30 484	F
B100	Biology A Modules B1, B2, B3	2014	1.37	3.08	0.41	15 694	F
B101	Biology A Modules B4, B5, B6	2012	2.34	7.84	0.37	1105	F
S108	Science Modules B1, C1, P1	2012	1.56	3.66	0.35	10 765	H
M108	Mathematics B	2012	1.60	3.80	0.34	18 784	F
S100	Science A Modules B1, C1, P1	2013	1.24	3.29	0.34	6730	F
B103	Biology A: Unit 1 (B1, B2, B3)	2010	1.53	3.85	0.33	4042	F
M108	Mathematics B	2013	1.99	4.86	0.30	27 972	F
S101	Science A Modules B2, C2, P2	2013	1.26	2.88	0.30	10 627	H
C107	Chemistry B: Unit 2 (C4, C5, C6)	2012	0.99	2.40	0.30	1304	F
S100	Science A Modules B1, C1, P1	2013	0.77	1.79	0.30	7982	H
S100	Science A Modules B1, C1, P1	2012	1.64	3.93	0.28	6862	F
P108	Physics Modules P4, P5, P6	2015	0.63	1.35	0.28	421	F
S108	Science Modules B1, C1, P1	2014	0.92	2.10	0.28	30 137	H
S109	Science Modules B2, C2, P2	2013	1.31	3.37	0.27	25 097	F
S104	Science A: Unit 2 (B2, C2, P2)	2011	1.74	4.47	0.26	17 360	F

Note: All 29 of these components were identified as outliers using Tukey's upper fence. The assessment codes were created by the researcher and are not the real codes.