ofqual

RESEARCH AND ANALYSIS

An exploration of the effect of speededness in a selection of GCSE examinations

Authors

- Qingping He
- Yasmine El Masri

How to cite this publication

He, Q. & El Masri, Y. H. (2025). *An exploration of the effect of speededness in a selection of GCSE examinations.* (Ofqual research report 25/7267/3). Ofqual. Retrieved from https://www.gov.uk/government/publications/an-exploration-of-the-effect-of-speededness-in-a-selection-of-gcse-examinations

Contents

Executive summary

Test speededness is defined as the extent to which test takers' scores depend on the rate at which they work as well as on the correctness of the responses (AERA, APA and NCME, 2014). Examinations for GCSE and A level in England do not explicitly include speed of responding as part of their assessment construct. In this study, we explored the effects of speededness in 181 GCSE examination papers from 6 subjects: mathematics, biology, chemistry, physics, combined science, and geography.

Our analysis was based around treating unanswered items at the end of the paper (that is, after the last item that had received a mark) as 'not reached'.

Omission rates of unreached items (estimated as the percentage of students who did not reach the item) varied across the papers from the different subjects. The omission rates for the biology, chemistry and physics papers were generally lower than those for the maths, combined science and geography papers. For most of the maths and combined science papers, less than 10% of items had an omission rate greater than or equal to 2%. Less than 5% of items had an omission rate greater than or equal to 2% for most of the biology, chemistry and physics papers, but more than 15% did for the majority of the geography papers. For the 5 tiered subjects, higher tier papers generally had lower item omission rates than foundation tier papers. The maximum item omission rates (the omission rates of the last items in the papers) varied from 3.3 to 51.5% (mean 20.8%) for mathematics, 1.8 to 25.1% (mean 7.4%) for biology, 1.0 to 8.4% (mean 4.3%) for chemistry, 0.5 to 15.8% (mean 4.7%) for physics, 2.5 to 58.1% (mean 14.5%) for combined science, and 5.8 to 15.4% (mean 9.7%) for geography, respectively. There was therefore substantial variation in the apparent degree of speededness across the papers from the different subjects.

There was considerable variability in item omission rate between the different population groups in the papers analysed. Students with special education needs (SEN) and those eligible for free school meals (FSMs) had the highest omission rates. Female students generally had lower item omission rates than male students.

The partial credit Rasch model was used to study the impact of possible speededness on students' performance and the results were used to classify students taking an exam paper into 3 groups: the most-speeded group, the less-speeded group and the least-speeded group. These accounted for 10%, 15% and 75% of the total number of students respectively. Multinomial logistic regression was used to examine the relationship between classification and students' background characteristics.

Based on the partial credit Rasch model, for the least-speeded groups, the mean of differences between the expected paper scores based on treating scores on

unreached items as missing and the expected scores based on treating scores on unreached items as zeros was close to zero for all the 181 papers analysed. For the less-speeded groups, for most of the papers, the mean of differences between the expected paper scores was less than 0.5% of the maximum paper marks. For the most-speeded groups, the mean of differences in expected paper scores ranged from about 0.62 to 4.52% (mean 1.54%) of maximum paper marks for mathematics, 0.46 to 3.28% (mean 1.22%) for biology, 0.77 to 4.21% (mean 1.61%) for chemistry, 0.52 to 3.71% (mean 1.39%) for physics, 0.17 to 7.16% (mean 2.00%) for combined science, and 1.93 to 5.89% (mean 3.66%) for geography, respectively.

Multinomial logistic regression analysis indicated that, for all the papers investigated, the probability of students being classified into the least-speeded group increased with ability[1], whereas that of being classified into the most-speeded group decreased (other things being equal). It appears that boys had higher probabilities of being classified into the most- and less-speeded groups than girls (other things being equal). For some of the papers, model coefficients for eligibility for FSM and SEN status were significantly different from zero, with those eligible for FSM and with SEN having higher probabilities of being classified into the most-speeded groups than those not eligible for FSM or without SEN (other things being equal). For first language and ethnicity, model coefficients for the majority of the papers were not significantly different from zero, suggesting similar impact from speededness with respect to these variables.

The main limitation of this study is that some of the assumptions made when estimating item omission rate and applying the Rasch model might not have been met fully by the data analysed. This needs to be borne in mind when considering the findings presented in this report. Further, speededness is only one of the many factors that could have contributed to the observed response patterns for the end-of-test items in the examination papers analysed in this study. Factors not considered include changes in topic areas, the level of demand or item difficulty, motivation, effort, fatigue, and others. Another important limitation of this study is the inclusion of students in the analysis who had received extra time when taking the papers. This is likely to have resulted in slightly underestimating the true levels of the effect of possible speededness in the papers and affected the relationships between classification and students' background characteristics to some extent.

[1] In this report, we follow the convention in psychometrics of using 'ability' to refer to the level of attainment in relation to the construct being assessed. It does not refer to IQ or SEN status.

1. Introduction

High-stakes tests and examinations, such as GCSEs and A levels used in England, are normally administered with time constraints for logistical reasons. There can be situations where some examinees cannot complete the test within the specified time limit. This can affect their performance on items located towards the end of the test, which is often characterised by either leaving these items unanswered or guessing. It is however to be noted that unusual responses to end-of-test items can result from a range of other possible causes, including, for example, increased level of demand or difficulty of end-of-test items, reduced level of effort or motivation to answer questions by examinees, fatigue, and others.

The effects of time constraints on test performance are referred to as test speededness effects (see Evans and Redly, 1972; Bolt, Cohen and Wollack 2002). Test speededness is defined as the extent to which test takers' scores depend on the rate at which work is performed as well as on the correctness of the responses (AERA, APA and NCME, 2024; also see Holmes (2025) for an in-depth discussion of the definitions of test speededness, factors affecting speededness and its impact on students' performance and potential implications). As suggested by Bolt et al. (2002) (also see Lu and Sireci, 2007; Hong and Cheng, 2019), speededness effects are often detrimental to the intended functioning of the test or examination, when speed of responding is not part of the construct intended to be measured by the test. Therefore, a high level of speededness could introduce construct-irrelevant variance into test scores and compromise the validity and fairness of the test.

If unreached or omitted items at the end of the test due to speededness are scored as zero then they might appear to be more difficult than they really are (see Bolt et al., 2002; Jin and Wang, 2014). However, if they are treated as "not attempted" then they might appear easier than they really are (for example if ability is correlated with speed of working and only the better students attempt them). When running out of time, examinees may also provide incorrect answers to end-of-test items (for example, random selection of answer options in multiple-choice questions), which will also make these items appear more difficult than they actually are.

There has been little research on exam speededness in the context of GCSE and A level examinations. Wheadon (2011) employed a mixture Rasch model for dichotomous items to investigate the effects of speededness in GCSE mathematics and science foundation papers. Results from this study suggested that a substantial proportion of students appeared to have run out of time in mathematics examinations. For one of the papers, nearly half of the students had an unexpectedly low probability of answering the last 5 questions correctly. In contrast, speededness did not appear to have affected students' performance on the science papers. More recently, Walland (2024) looked at speededness in 340 GCSE written papers from

2009 to 2016 (pre-reform qualifications) based on average percentage marks lost from the longest string of unanswered items at the end of the paper of the students (that is average maximum available marks associated with unreached items expressed as percentage of the maximum available paper mark). She suggested that most exam papers she analysed were unlikely to be speeded.

The aim of the study reported here was to explore the effects of possible speededness in 181 reformed GCSE examination papers from mathematics, biology, chemistry, physics, combined science, and geography, administered in 2017, 2018 and 2019[2]. The focus was particularly on the extent of possible impact of speededness on exam performance and how students with different background characteristics were likely to be affected.

2. Methodology

2.1 Data

Students' item-level data from 181 exam papers from 6 GCSE subjects from 2017 to 2019 was analysed in this study (papers taken by less than 1,000 students were excluded from the analysis). These were mathematics, biology, chemistry, physics, combined science, and geography. Each subject had a number of specifications[3] across the exam boards that offered the qualification. Except for the specifications in geography, each specification for the other subjects had 2 tiers (a higher tier targeted at grades 4 to 9 and a foundation tier targeted at grades 1 to 5) and each tier had a number of exam papers. Table 1 lists the number of papers from the individual specifications investigated, the number of items and maximum paper marks, and the years of administration that we investigated. The number of items in the papers varied from 17 in a geography paper to 63 in a combined science foundation paper. All questions in all the papers analysed were compulsory. The maximum number of marks of the items varied from 3 for a higher tier chemistry paper to 12 for a geography paper.

[2] The reformed GCSE mathematics was examined for the first time in 2017, and the other subjects in this study in 2018.

[3] Qualification specifications set out the knowledge, skills and understanding that must be studied, as well as the assessment method. A GCSE or A level specification for a particular subject will be based on the subject content published by the Department for Education and the assessment arrangements published by Ofqual. There may be some differences in how each individual specification in a particular subject approaches the detail of the content and assessment, although each will attribute the same proportion of marks to the identified key abilities.

Table 1 Number of items, maximum marks of the items, and maximum paper marks for the 181 papers analysed

Subject	Spec	Tier	Year	No of papers	No of items	Max item mark	Max paper mark
Maths	SP1	H	2017-19	9	30-38	4,5,6	80
		F	2017-19	9	34-45	4,5,6	80
	SP2	H	2017-19	9	30-39	5,6,7	100
		F	2017-19	9	41-50	5,6	100
Biology	SP1	H	2018, 19	4	46-48	6	100
		F	2018, 19	4	53-60	6	100
	SP2	H	2018, 19	4	46-54	6	90
		F	2019	2	48,54	6	90
	SP3	H	2018, 19	4	34-56	4,6	90
Chemistry	SP1	H	2018, 19	4	46-48	6	100
		F	2018, 19	4	53-60	6	100
	SP2	H	2018, 19	4	46-54	6	90
		F	2019	2	51-57	6	90
	SP3	H	2018, 19	4	37-53	3,4,6	90
Physics	SP1	H	2018, 19	4	40-48	6	100
		F	2018, 19	3	52-62	6	100
	SP2	H	2018, 19	4	44-54	6	90
		F	2019	2	47,57	5,6	90
	SP3	H	2018, 19	4	33-46	4,6	90
Combined science	SP1	H	2018, 19	8	38-45	6	100
		F	2018, 19	8	58-63	6	100
	SP2	H	2018, 19	12	26-35	6	70
		F	2018, 19	12	31-46	6	70
	SP3	H	2018, 19	12	28-37	6	60
		F	2018, 19	12	30-40	6	60
	SP4	H	2018, 19	7	30-46	6	75,95
		F	2018, 19	8	34-49	6	75,95
Geography	SP1		2019	1	24	9	76
	SP2		2018, 19	6	18-28	8,12	60,80
	SP3		2018, 19	6	17-26	8,12	60,70

Students' demographic data was also used in the analysis. This includes sex, eligibility for free school meals (FSMs), status of special education needs (SEN), home first language indicator (whether or not English is used as an additional language at home (EAL)), and ethnicity, extracted from the National Pupil Database .

2.2 Methods of analysis

In this analysis, two assumptions are made. First, it is assumed that students answer questions in the same sequence as they appear in the exam. It is also assumed that the reason items are not answered at the end of a test is because students have run out of time and did not reach the items. The distribution of item omission rate (OR), calculated as the percentage of students who did not reach the item, can therefore provide some indication about the location of the item in the test where the effect of speededness was likely to have started manifesting. The potential effect of speededness on performance is explored through analysis of omission rates of

unreached items and application of the partial credit Rasch model. This is detailed below.

Unattempted and unreached items

When GCSE and A level exams are marked, marks are given to students for attempted questions if their answer is correct or partially correct. These are the “valid” marks. In contrast, for unattempted (omitted) questions, special symbols or numbers are given (referred to as “invalid” marks). Unattempted questions could be located anywhere in the exam paper as students may skip questions because they forgot to answer them, they did not see them, they found them too hard or, for questions located at the end of a paper, they did not reach them.

In this study, all the questions before the last question with a valid mark were treated as reached and attempted even if, in reality, students might have intentionally skipped some of them (a similar assumption was made by other researchers – see, for example, Borter, Schlegel and Troche, 2023). Those questions were given a score of zero because we assumed that they were too difficult for the students to answer correctly had they attempted them. In contrast, we treated omitted questions after the last question with a valid mark as unreached and coded them as missing data because we assumed that students had run out of time.

It is important to note that these assumptions may not reflect reality. As indicated before, there can be other reasons for unattempted questions in a paper. It is also to be noted that the datasets analysed include data for students who had received extra time when taking the exams. It was not possible to identify these students and to exclude them from the analysis. As a result, the omission rates calculated for unreached items are likely to have been underestimated, although the size of underestimation is likely to be small.

To assess the impact of speededness on performance, we calculated the mean score for unreached items twice. In the first instance, the means were estimated with all unreached questions being treated as missing. In the second instance, the means were estimated with unreached questions being treated as incorrect and hence scored as zero, which is the normal practice when marking GCSE and A level exams.

Application of the partial credit Rasch model

A variety of methods have been used to study test speededness (see Lu and Sireci, 2007; Jin and Wang, 2014; Cintron, 2021; Yu and Cheng, 2022). Lu and Sireci (2007) described several methods that can be used to detect and evaluate the level of speededness in tests composed of multiple-choice or dichotomously scored items (that is, items with 2 score categories). However, these methods are not suitable for

this study as the majority of the items here are polytomous (items with more than 2 score categories). Here we have used the partial credit Rasch model to investigate the effect of speededness in the exam papers.

The partial credit Rasch model (PCM) (see Masters, 1982; Wright and Masters, 1982) is a mathematical function used to describe the probability $P_{n,i}(\theta_n, x)$ of an examinee n with ability (latent trait) θ_n scoring x on an item i:

$$P_{n,i}(\theta_n, x) = \frac{\exp \sum_{k=0}^{x}(\theta_n - \delta_{ik})}{\sum_{l=0}^{m} \exp[\sum_{k=0}^{l}(\theta_n - \delta_{ik})]} \quad (1)$$

where m is the maximum available score on the item and δ_{ik} is the location of the kth score category (k>0) on the latent trait continuum (also termed category threshold or difficulty, see Andrich, 2015).

In the present study, for unreached items, students' scores were treated as missing, and item parameters δ_{ik} and students' abilities A_n (and the expected score on the paper $S_{n,exp}$, calculated as the sum of the expected scores on individual items) were then estimated. The item parameter values were then fixed and the scores on unreached items were treated as zero, and students' abilities A'_n (and the expected score on the paper $S'_{n,exp}$) were re-estimated. The difference between the estimates of student abilities with unreached responses treated as incorrect A'_n ($S'_{n,exp}$) and the estimates with unreached responses treated as missing A_n ($S_{n,exp}$) would represent the impact of speededness on performance. Higher values would indicate a larger effect.

The differences in expected paper scores $S_{n,exp} - S'_{n,exp}$ were used to classify students into 3 speeded groups: the least-speeded group (consisting of students with the smallest difference values, accounting for 75% of the total candidature), the less-speeded group (accounting for 15% of the total candidature), and the most-speeded group (consisting of students with the largest values, accounting for 10% of the total candidature). Although we note that these percentages were arbitrarily set, they in some way reflect the fact that, for the majority of the papers analysed here, the majority of the students were able to complete the papers. This suggests that the effect from speededness in these papers is likely to be small for the majority of the students and only substantial for a small proportion of them.

The R package *TAM* which implements the PCM and marginal maximum likelihood (MML) estimation method was used for conducting the Rasch analysis (see Robitzsch, Kiefer and Wu, 2022).

It is noted that while this report was being written, Walland (2024) published her research which investigated the effect of speededness in 340 GCSE written papers from 2009 to 2016 offered by one of the exam boards. This used average percentage marks lost from the longest string of unanswered items at the end of the paper of the students (that is average maximum available marks associated with

unreached items expressed as percentage of the maximum available paper mark). Whereas this approach might be simple and effective in identifying speeded exams, the Rasch analysis approach adopted in our study could be used to investigate the impact of speededness on the performance of individual students. We have, however, also used Walland's approach to analyse the 181 GCSE exam papers studied here and included a brief discussion, in Appendix A, of the results obtained.

Relationship between speededness and students' characteristics

To examine how the classification of students into different speeded groups is affected by their background characteristics, the speeded groups were treated as a categorical variable. A multinomial logistic regression was used to model how the probabilities of students being classified into individual speeded groups are affected by their ability and other characteristics. For student n, let P_{nk} and P_{nR} be the probabilities of being classified into group k and the reference group R. Given a student's background characteristics, the logarithm of the ratio of the 2 probabilities is expressed as a linear function of the background variables:

$$ln\left(\frac{P_{nk}}{P_{nR}}\right) = \beta_{0k} + \beta_{1k}A_n + \beta_{2k}Sex_n + \beta_{3k}FSM_n + \beta_{4k}EAL_n + \beta_{5k}Ethn_n + \beta_{6k}SEN_n + \beta_{2ki}A_n \times Sex_n + \beta_{3ki}A_n \times FSM_n + \beta_{4ki}A_n \times EAL_n + \beta_{5ki}A_n \times Ethn_n + \beta_{6ki}A_n \times SEN_n \quad (2)$$

where:

A_n = student's ability estimated using the PCM (continuous);

Sex_n = sex (binary, "F" for females, "M" for males – the reference category);

FSM_n = eligibility for FSM (binary, 0 for ineligible – the reference category, 1 for eligible);

EAL_n = home first language indicator (binary, 0 for English – the reference category, 1 for others);

$Ethn_n$ = ethnicity (binary, 0 for white – the reference category, 1 for others);

SEN_n = SEN status (binary, 0 for non-SEN – the reference category, 1 for SEN students);

β_{0k} = intercept for group k;

β_{1k} to β_{6k} = model coefficients for group k associated with the various independent variables;

β_{2ki} to β_{6ki} = model coefficients for the interaction terms.

For the 5 categorical independent variables (population subgroups), groups not used as the reference group are referred to as the focal groups. The least-speeded group was used as the reference group (R) for the dependent variable in the analysis. A non-zero coefficient (excluding the intercept) would suggest varying impact of speededness on the corresponding independent variable or its interaction with ability. The R package *nnet* was used to conduct the analysis (see Ripley and Venables, 2021).

3. Results

This section presents the main findings of the study.

3.1 Distribution of item omission rate in papers

Figure 1 below shows the distributions of omission rate across the items in 3 higher tier papers (top) and 3 foundation tier papers (bottom) from a GCSE maths specification in 2019. All unattempted items were retained as omitted in the graphs on the left, while only unreached items were retained as omitted in the graphs on the right. For the graphs on the left, while there appears to be a clear trend showing increase in omission rate across the course of the papers, for 5 of the 6 papers, maximum omission rate was not associated with the last item in the papers. There was also considerable variability in omission rate between items in the same paper and between the papers. Item omission rates in the foundation papers were generally higher than in the higher tier papers; and there were more items in the foundation tier papers with noticeable level of omission than in the higher tier papers. For the foundation tier papers, some of the items in the middle also showed higher omission rate than later items. For the higher tier papers, the maximum omission rates were 27.6% for Paper 1, 15.2% for Paper 2 and 22.6% for Paper 3 respectively. For the 3 foundation tier papers, these were 24.7%, 31.8% and 44.4%. For the graphs on the right, item omission rate for the unreached items increases monotonically towards the end of the papers, with the last item having the highest omission rate. The maximum omission rates were 13.9% for Paper 1, 15.2% for Paper 2 and 11.1% for Paper 3 respectively for the higher tier papers. For the 3 foundation tier papers, these were 10.1%, 7.2% and 16.3%.

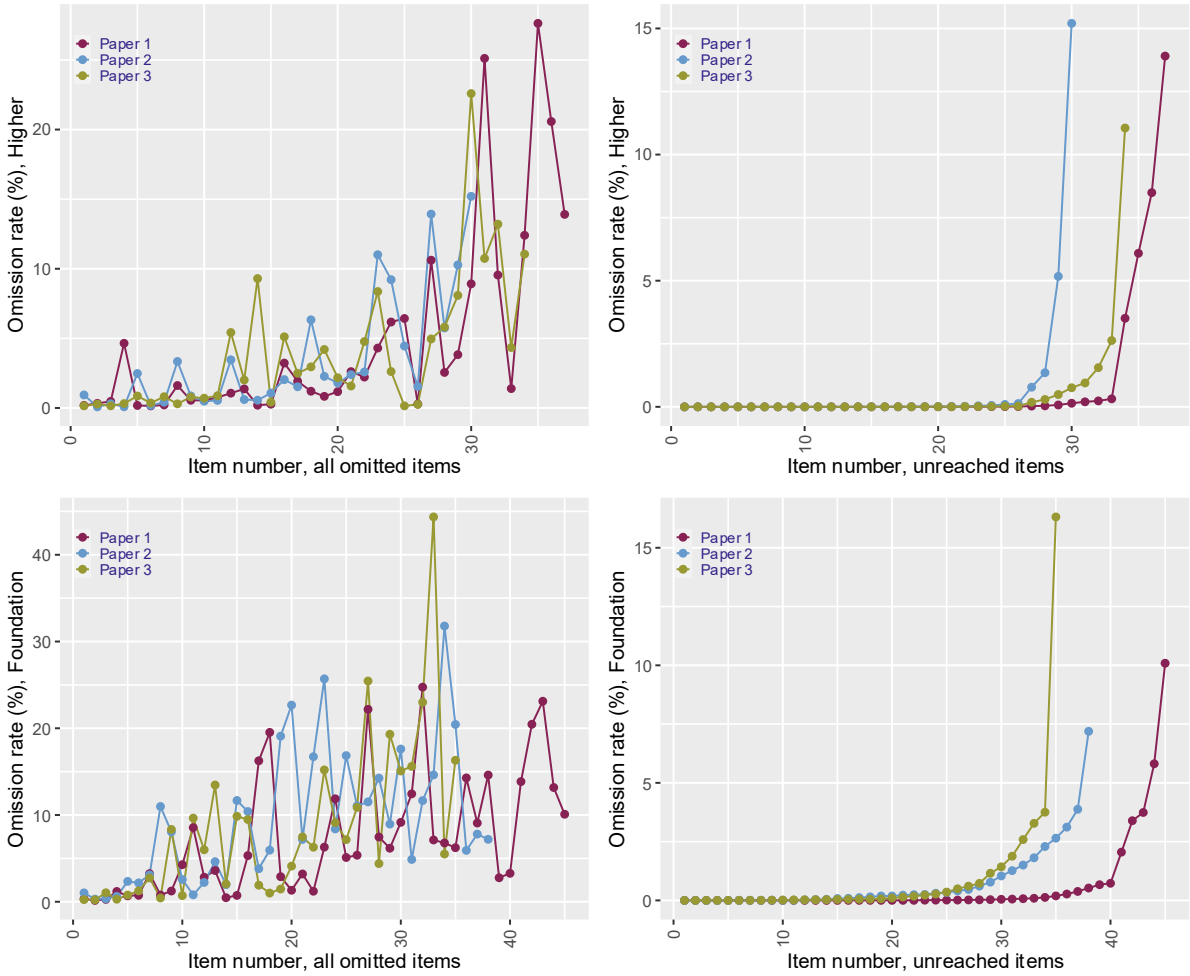

Figure 1 Distributions of omission rate across the items in 3 higher tier papers (top) and 3 foundation tier papers (bottom) from a GCSE maths specification in 2019. left: scores on all unattempted items were treated as missing; right: scores on unreached items were treated as missing

The possible location where speededness started in an exam paper and its potential impact on performance could be visualised graphically by examining the distributions of item omission rate like those shown in the graphs on the right in Figure 1. To see how the level of omission rate of unreached items varies across the 181 papers from the 6 different subjects (omission rate will be associated with unreached items only hereafter), Figure 2 shows distributions of the percentage of items in the papers with omission rate greater than or equal to 1%, 2% and 5% respectively. (Note: the x-axis in Figure 2 represents papers from individual subjects and the points on a vertical line in a graph are for the same paper. This also applies to Figures 5, 6, 8, 11, 12 and 14.) Table 2 shows the corresponding values of the range and mean for the different subjects.

For the 5 subjects with 2 tiers, the foundation tier papers generally showed a larger proportion of items with rates of omission above the cut-off points than higher tier papers. Maths and combined science papers had a larger proportion of items with omission rates above the cut-off points than the biology, chemistry and physics

papers. For most of the biology, chemistry and physics papers, the percentage of items with omission rate greater than or equal to 2% was less than 5%, whereas this was above 5% for most of the maths and combined science papers. Some of the maths and combined science foundation papers had over 10% of their items with omission rate equal to or greater than 5%.

The majority of the geography papers had more than 15% of their items with omission rate equal to or greater than 2%. Two of the papers had more than 10% of their items with omission rate equal to or greater than 5%.

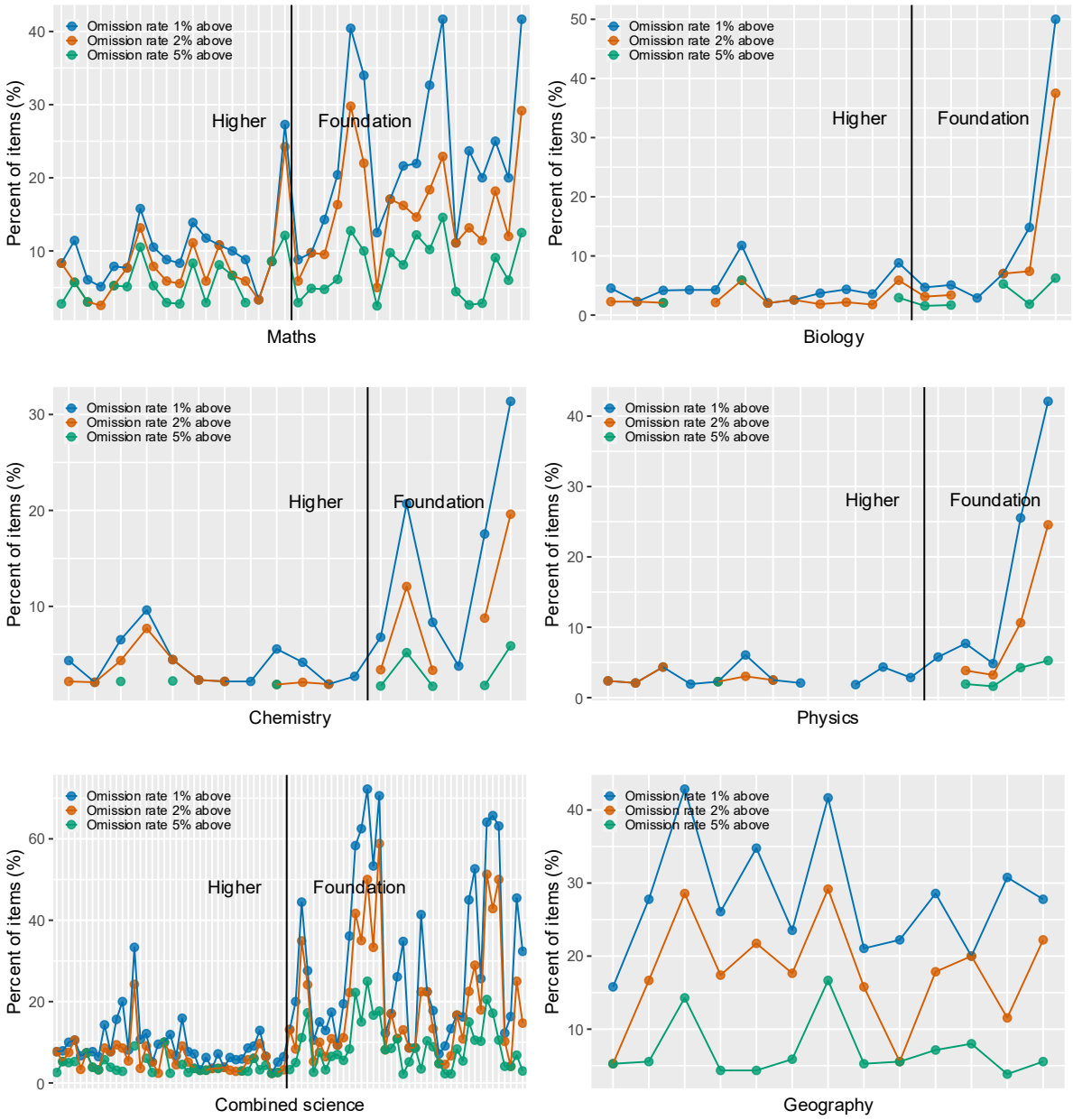

Figure 2 Distributions of percentage of items with omission rate equal to or greater than 1%, 2%, and 5% respectively for the 181 papers analysed. Note: the x-axis represents papers from individual subjects and the points on a vertical line in a graph are for the same paper. The black vertical line divides the two tiers for the 5 tiered subjects. This also applies to Figures 5, 6, 8, 11, 12 and 14.

Table 2 Range and mean of estimated percentages of items in the 181 papers with omission rate greater than or equal to 1%, 2% and 5% respectively

Subject	Tier	N	Range (mean) (OR ≥ 1%)	Range (mean) (OR ≥ 2%)	Range (mean) (OR ≥ 5%)
Maths	H	18	3.33-27.27 (10.25)	2.56-24.24 (7.87)	0.00-12.12 (5.17)
	F	18	8.82-41.67 (23.15)	5.00-29.79 (15.70)	2.50-14.58 (7.57)
Biology	H	12	2.04-11.76 (4.69)	0.00-5.88 (2.58)	0.00-5.88 (0.91)
	F	6	2.90-50.00 (14.08)	0.00-37.50 (9.74)	0.00-6.25 (2.77)
Chemistry	H	12	1.89-9.62 (4.00)	0.00-7.69 (2.59)	0.00-2.22 (0.52)
	F	6	3.77-31.37 (14.75)	0.00-19.61 (7.86)	0.00-5.88 (2.70)
Physics	H	12	0.00-6.06 (2.73)	0.00-4.35 (1.38)	0.00-2.27 (0.19)
	F	5	4.84-42.11 (17.19)	0.00-24.56 (8.45)	0.00-5.26 (2.61)
Combined science	H	39	2.38-33.33 (9.05)	2.38-24.24 (6.01)	0.00-10.00 (3.16)
	F	40	7.14-72.22 (30.42)	4.08-58.82 (20.25)	2.17-25.00 (9.12)
Geography		13	15.79-42.86 (27.91)	5.26-29.17 (17.65)	3.85-16.67 (7.05)

Figure 3 further shows the distributions of omission rates of the last items in the 181 papers analysed. These are the maximum item omission rates and vary from 3.3 to 51.5% (mean 20.8%) for mathematics, 1.8 to 25.1% (mean 7.4%) for biology, 1.0 to 8.4% (mean 4.3%) for chemistry, 0.5 to 15.8% (mean 4.7%) for physics, 2.5 to 58.1% (mean 14.5%) for combined science, and 5.8 to 15.4% (mean 9.7%) for geography, respectively. The maximum omission rates for the biology, chemistry and physics papers are generally lower than those for the maths, combined science and geography papers. Except for chemistry, foundation tier papers generally have higher maximum omission rate than higher tier papers (see Table 3). The omission rate of the last item in a paper or the maximum omission rate represents the proportion of students who did not complete the paper.

Table 3 Range and mean of estimated maximum item omission rates for the 181 papers from the different subjects

Subject	Tier	N	Max omission rate range (mean) (%)
Maths	H	18	3.27-32.02 (16.37)
	F	18	5.54-51.41 (25.16)
Biology	H	12	1.96-12.64 (4.12)
	F	6	1.83-25.11 (14.19)
Chemistry	H	12	1.13-8.36 (4.35)
	F	6	1.91-8.88 (6.66)
Physics	H	12	0.48-9.73 (3.13)
	F	5	1.79-15.76 (8.68)
Combined science	H	39	2.49-29.35 (9.34)
	F	40	7.55-58.06 (19.51)
Geography		13	5.79-15.38 (9.72)

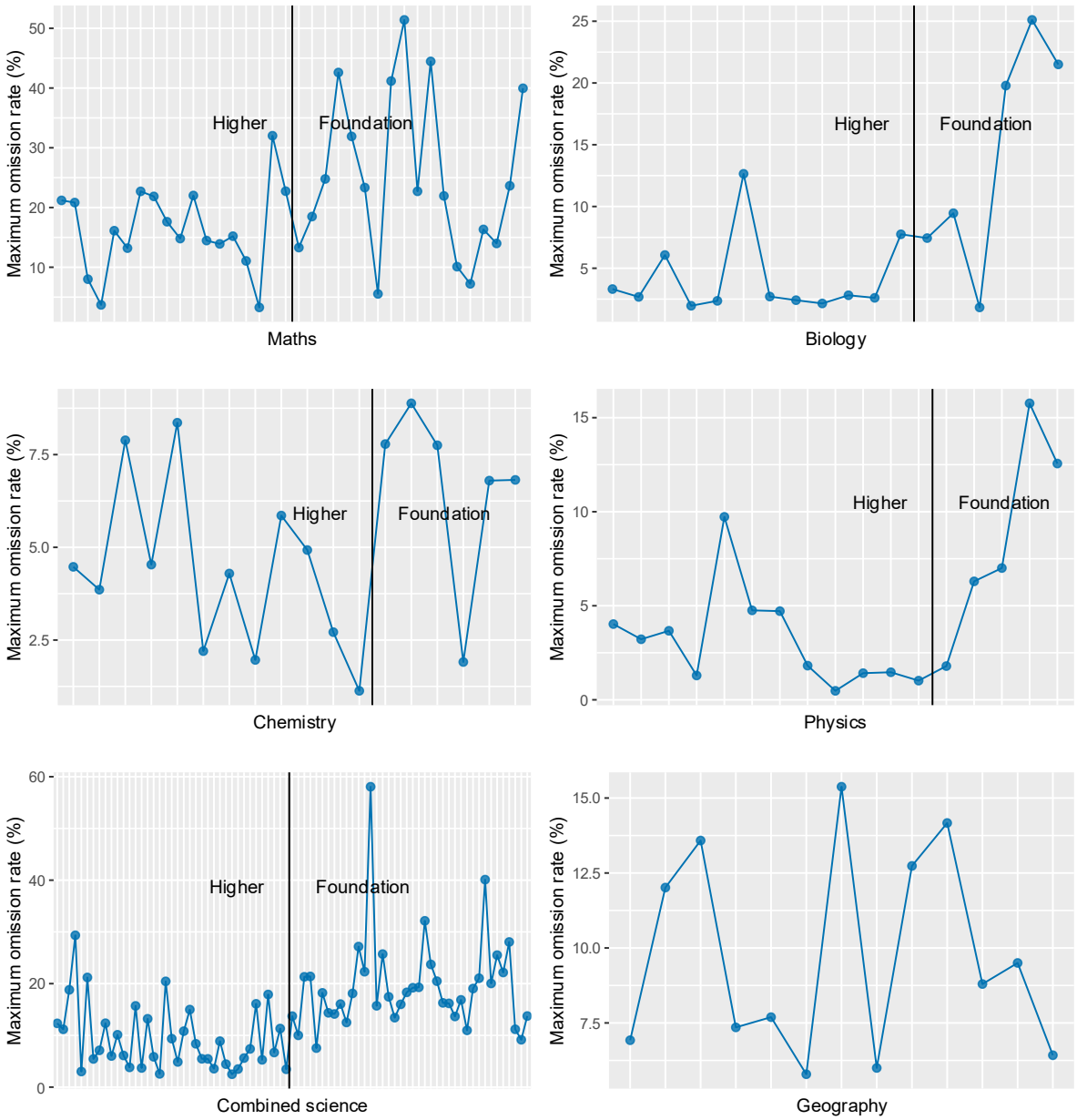

Figure 3 Distributions of estimated maximum item omission rate for the 181 papers analysed

3.2 Variation of item omission rate between population subgroups

The variability of item omission rate between population subgroups in the papers was also explored. As an example, Figure 4 shows the distribution of omission rate for items in the 3 higher tier maths papers and the 3 foundation tier papers discussed earlier (see also Figure 1), for individual subgroups. For the 6 papers, students with SEN and those eligible for free school meals had the highest omission rates whereas

non-white students and students with English as an additional language had the lowest omission rates. Male students had higher omission rates than female students. Note that except for sex, the number of students in the focal group is substantially smaller than in the reference group.

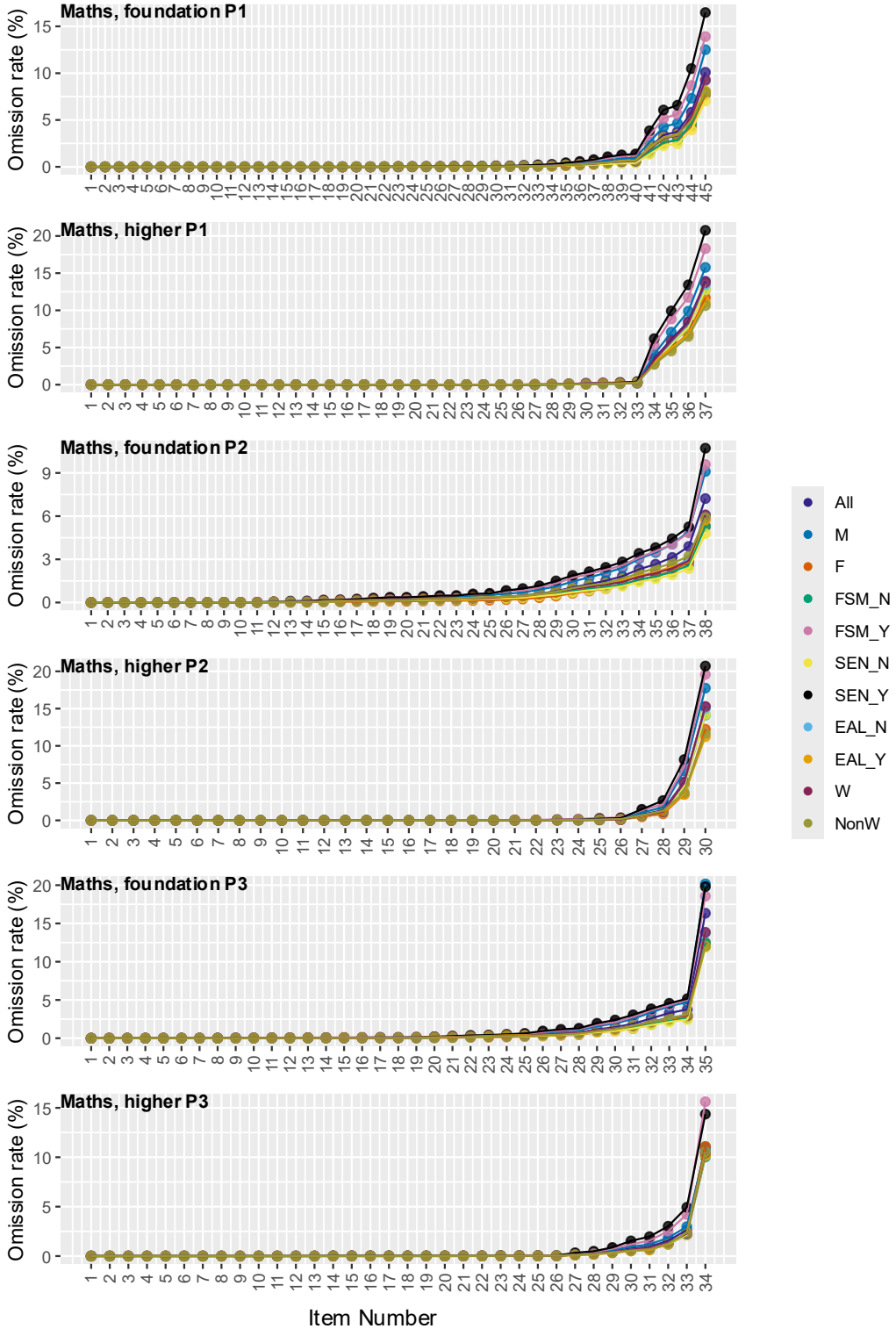

Figure 4 Distributions of estimated item omission rate across the items in the 3 higher tier maths papers (left) and the 3 foundation tier papers (right) for the different population subgroups

Figure 5 shows the distributions of the maximum item omission rate for items in the 181 papers for the different subgroups. There is substantial variability in omission rate between the different subgroups. In general, students with SEN and those eligible for FSMs had the highest omission rates, while female students had the lowest omission rate.

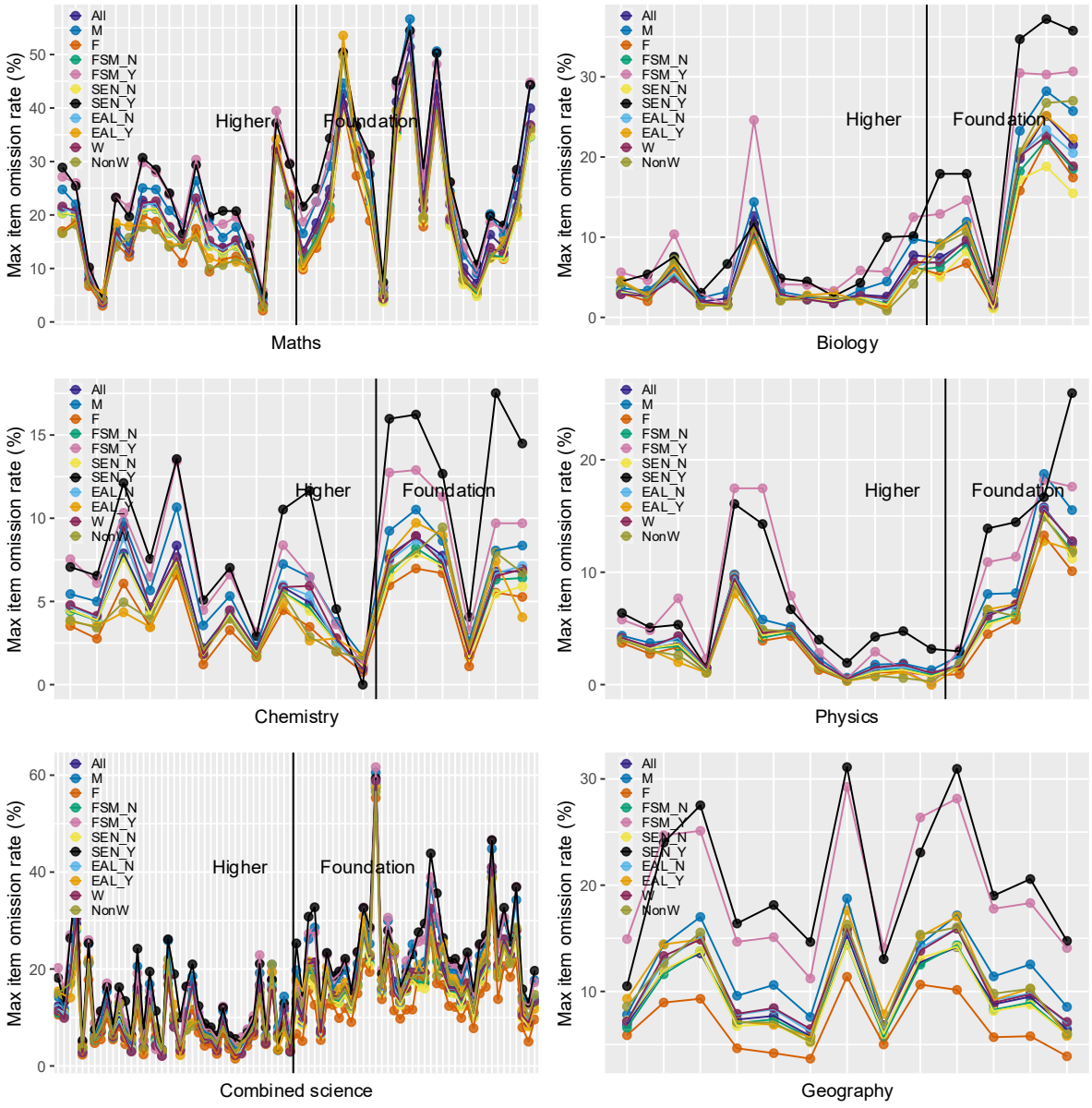

Figure 5 Distributions of maximum item omission rate across the 181 papers for different population subgroups

To have a clearer picture of the differences in omission rates between the subgroups in the different papers, Figure 6 shows the distributions of the differences in maximum item omission rate between the focal groups and the reference groups. Positive differences indicate that the omission rate is higher for the focal group than the reference group, whereas negative values indicate that it is lower. The differences are larger for maths, combined science, and geography than for biology, chemistry and physics. For some of the maths and combined science papers,

particularly those from the foundation tiers, the maximum omission rates for female students were over 10% lower than those for male students. For a small proportion of the maths and combined science papers and some of the geography papers, the omission rates for students with SEN were over 10% higher than those without. For some of the geography papers, the maximum omission rates for students eligible for FSMs were over 10% higher than those not eligible.

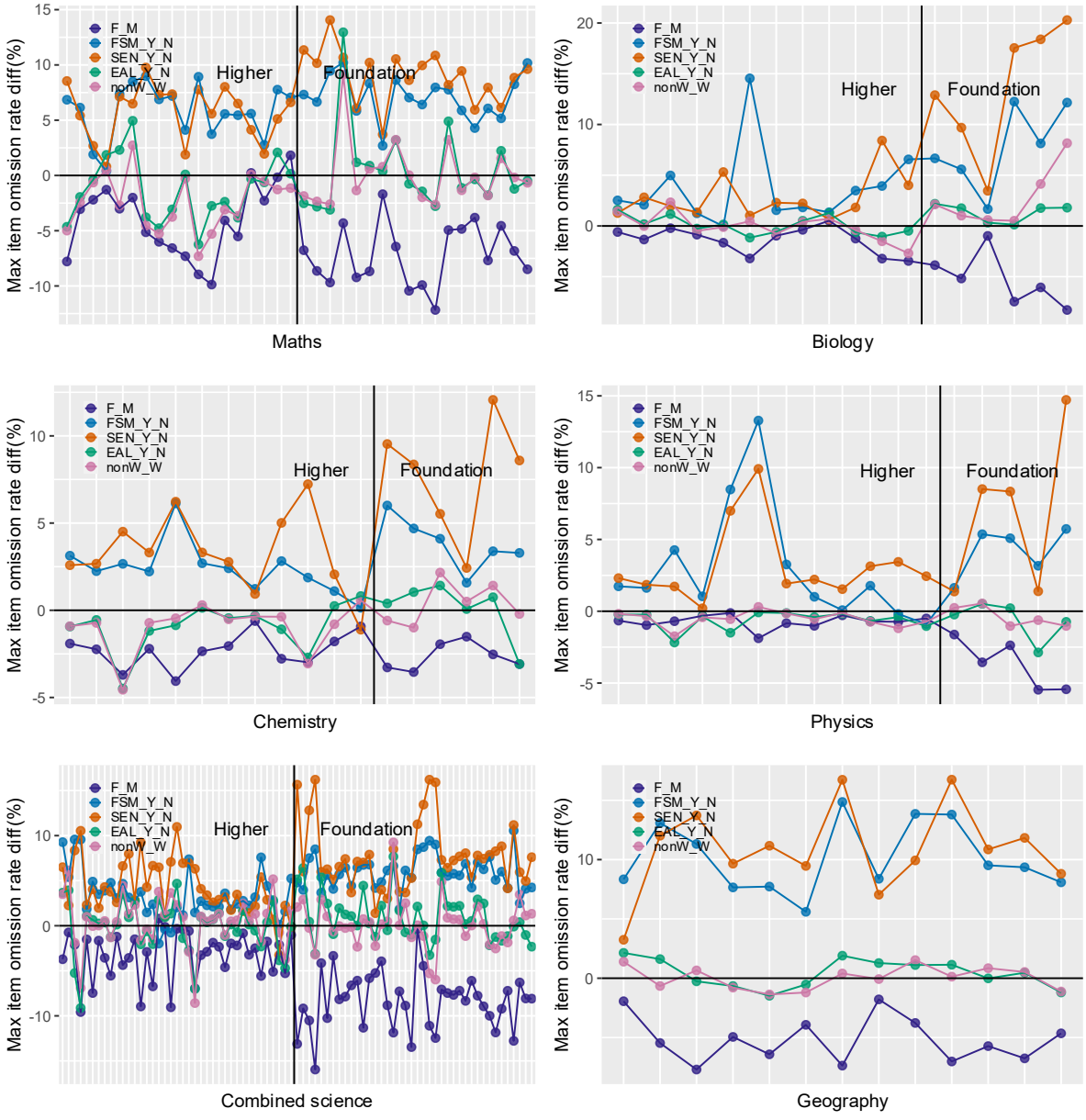

Figure 6 Distributions of differences in maximum item omission rate between the focal groups and the reference groups across the 181 papers analysed

For ethnicity and home first language, the differences in maximum item omission rate between the focal groups and the references groups were generally very small for the 3 separate science subjects and geography. For maths and combined science, the differences were still noticeably smaller than for the other group variables, although they varied considerably between the papers

3.3 Marks associated with unreached items

Figure 7 shows the distributions of estimated item mean score (expressed as the percentage of maximum available mark) across the items in the 3 higher tier maths papers and the 3 foundation tier papers. For all the 6 papers, item mean score generally decreases towards the end of the paper, although there is considerable variability between items within the same paper and between the papers. Similar patterns are observed for all the other papers investigated. For items near the end of the papers, the mean scores are slightly higher when treating scores on unreached items as missing than as zeros (for earlier items the dots and lines overlap each other in the figure).

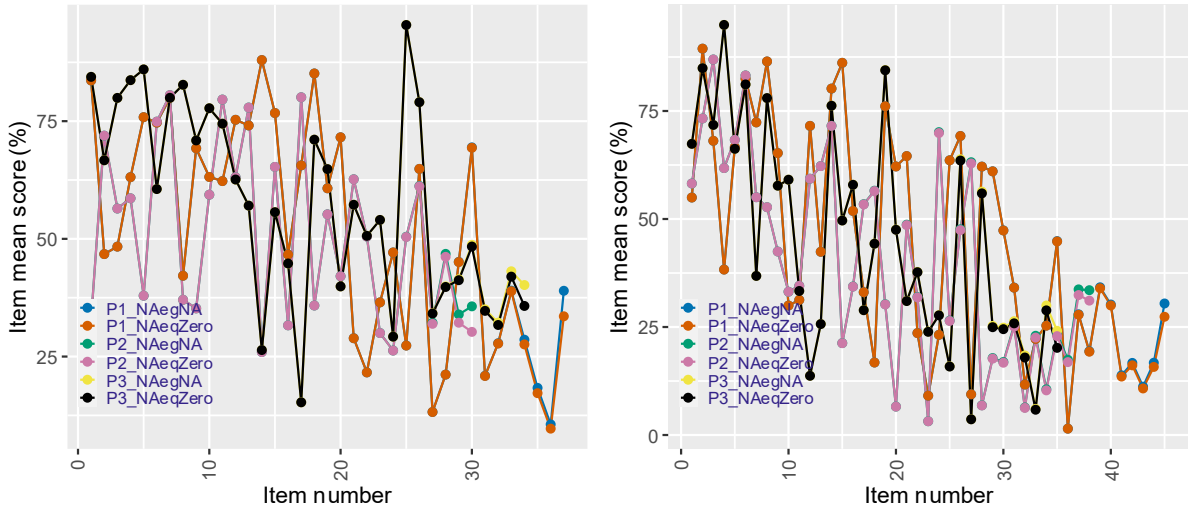

Figure 7 Distributions of estimated item mean score (expressed as the percentage of maximum available mark) across the items in 3 higher tier maths papers (left) and the 3 foundation tier papers (right). In the legend keys, “NAeqNA” represents treating scores on unreached items as missing and “NAeqZero” treating scores on unreached items as zeros

To see the extent of marks in the papers associated with unreached items further, Figure 8 shows the percentages of total paper marks that are associated with items with omission rate greater than or equal to 1%, 2% and 5% respectively. Table 4 shows the corresponding values of range and mean for the different subjects. Substantial variability exists in the marks between papers from the same subject and between the different subjects.

For maths and combined science, percentages of marks associated with unreached items in the foundation tier papers were considerably higher than in the higher tier papers. For maths, percentages of marks associated with unreached items with omission rate greater than or equal to 1%, 2% and 5% for the 3 higher tier papers varied from 4.0%, 2.0%, and 0% of the maximum paper marks respectively to 23.0%, 19.0% and 12.5%, whereas these varied from 8.8%, 5.0%, and 2.5% to 50.0%, 38.0% and 19.0% for the foundation tier papers. For combined science,

these varied from 3.2%, 2.1%, and 0% respectively to 40.0%, 25.0%, and 11.7% for the higher tier papers and from 9.0%, 5.3%, and 2.9% to 83.3%, 70.0% and 33.3% for the foundation tier papers.

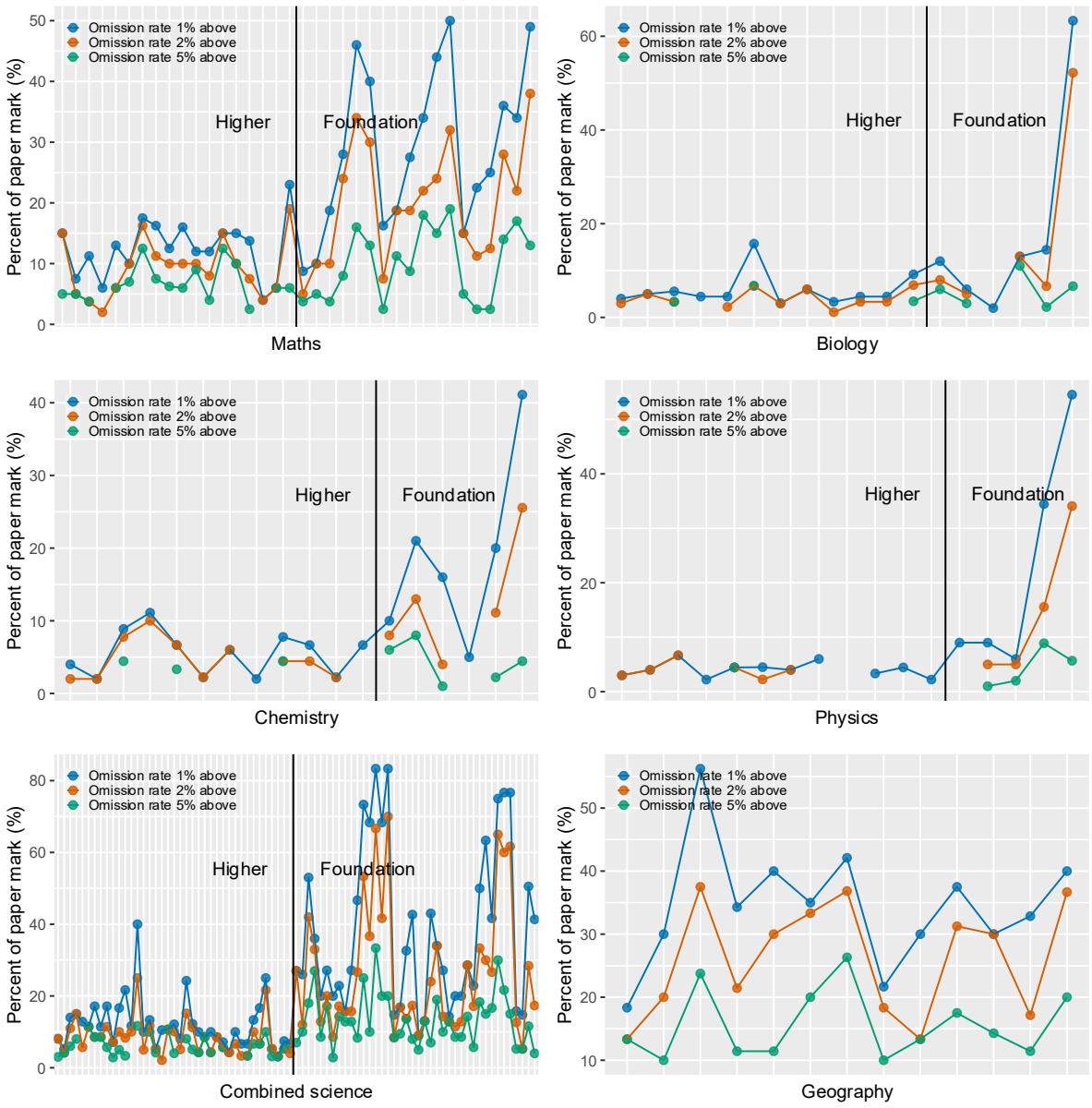

Figure 8 Distributions of percentage of total paper marks associated with items with omission rate ≥1%, ≥2% and ≥5%, respectively, in the 181 papers analysed

Table 4 Range and mean of marks associated with items with omission rate greater than or equal to 1%, 2% and 5%, respectively, expressed as percentages of maximum paper marks for the 181 papers

Subject	Tier	N	Range (mean) (OR ≥ 1%) (%)	Range (mean) (OR ≥ 2%) (%)	Range (mean) (OR ≥ 5%) (%)
Maths	H	18	4.00-23.00 (12.54)	2.00-19.00 (9.38)	0.00-12.50 (6.06)
	F	18	8.75-50.00 (29.08)	5.00-38.00 (20.15)	2.50-19.00 (9.89)
Biology	H	12	3.00-15.73 (5.80)	0.00-6.90 (3.66)	0.00-6.74 (1.13)
	F	6	2.00-63.33 (18.46)	0.00-52.22 (14.15)	0.00-11.00 (4.81)
Chemistry	H	12	2.00-11.11 (5.52)	0.00-10.00 (3.98)	0.00-4.44 (1.02)
	F	6	5.00-41.11 (18.85)	0.00-25.56 (10.28)	0.00-8.00 (3.61)
Physics	H	12	0.00-6.67 (3.74)	0.00-6.67 (2.03)	0.00-4.44 (0.37)
	F	5	6.00-54.55 (22.60)	0.00-34.09 (11.93)	0.00-8.89 (3.51)
Combined science	H	39	3.16-40.00 (11.93)	2.11-25.00 (8.35)	0.00-11.67 (4.66)
	F	40	9.00-83.33 (38.57)	5.26-70.00 (26.82)	2.86-33.33 (13.34)
Geography		13	18.33-56.25 (34.46)	13.33-37.50 (26.09)	10.00-26.32 (15.60)

For most of the higher tier papers from biology, chemistry, and physics, less than 10% of the maximum paper marks were associated with items with omission rate greater than or equal to 1% and less than 5% of the paper marks were associated with items with omission rate greater than or equal to 5%. For the foundation tier papers, there were one or two papers from each of the subjects with over 40% of the paper marks associated with items with omission rate greater than or equal to 1%, over 25% of the paper marks associated with items with omission rate greater than or equal to 2%, and about 10% of the paper marks associated with items with omission rate greater than or equal to 5%.

For the geography papers, marks associated with unreached items were relatively high. Percentages of maximum paper marks associated with items with omission rate greater than or equal to 1%, 2% and 5% varied from 18.3%, 13.3% and 10.0% respectively to 56.3%, 37.5% and 26.3%.

Overall, the findings discussed previously suggest that there was a degree of speededness in some of the papers analysed here. We have also used the Walland (2024) approach to analyse the papers investigated in this study and a brief discussion of the results obtained is included in Appendix A. This also suggests a degree of speededness in some of the papers, particularly in some of the maths and combined science foundation tier papers.

3.4 The impact of speededness on performance based on the partial credit Rasch model

This section looks at the impact of speededness on students' performance based on analysis using the Rasch partial credit model (PCM).

Figure 9 shows the distributions of Rasch item difficulties (calculated as the means of category thresholds) for the 3 higher tier maths papers and the 3 foundation tier papers discussed earlier, based on treating scores on unreached items as missing. For each paper in Figure 9, items are arranged in the same order as their sequence in the paper. Item difficulty generally increases towards the end of the papers, although there is variability between items in the same paper and between the different papers.

When comparing the distributions in Figure 9 and those in Figure 7, it should be borne in mind that Figure 7 shows mean facility scores which are measures of easiness while Figure 9 shows distributions of item difficulty estimates. That is, items with relatively low mean scores in Figure 7 will have relatively high values of difficulty in Figure 9. The distributions of item difficulty estimates and mean facility scores shown in Figures 7 and 9 are broadly similar in terms of trend of estimated item difficulty in the papers. Similar patterns are observed for all the other papers analysed.

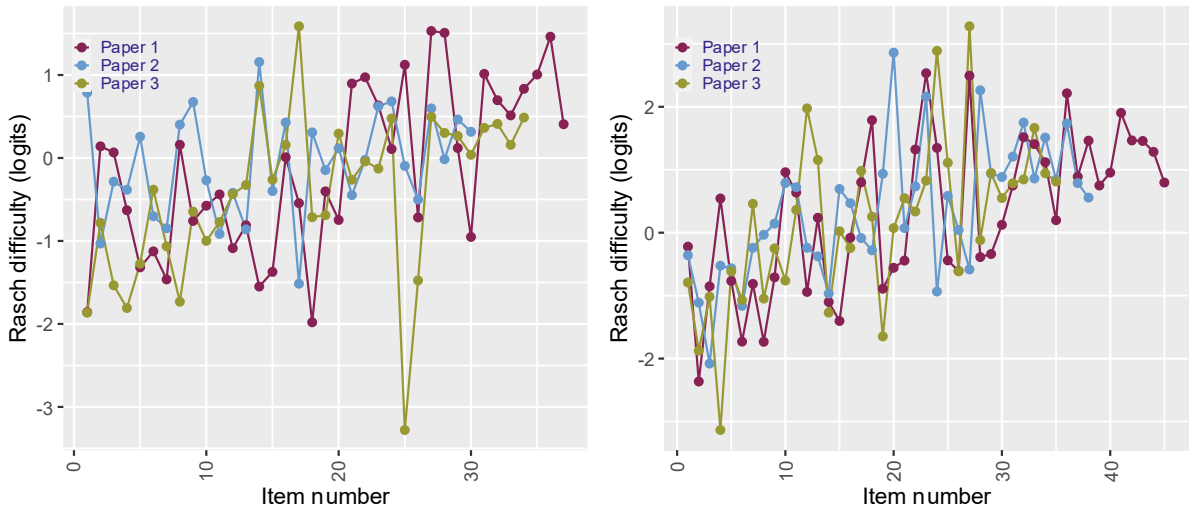

Figure 9 Distributions of item difficulties for items in the 3 higher tier maths papers (left) and 3 foundation tier papers (right) derived based on the partial credit Rasch model (scores on unreached items were treated as missing)

Impact of speededness on scores at paper level

Figure 10 compares students' expected paper scores $S_{n,exp}$ based on treating scores on unreached items as missing (y-axis) against the expected scores $S'_{n,exp}$ based on treating scores on unreached items as zeros (x-axis) for the 6 maths papers. Similar relationships between the 2 sets of expected scores are seen for all the other papers. As is clear from Figure 10, for some of the students, $S_{n,exp}$ are considerably higher than $S'_{n,exp}$, particularly those in the middle of the ability range. These students were likely to have been affected more by speededness than students with high or low abilities.

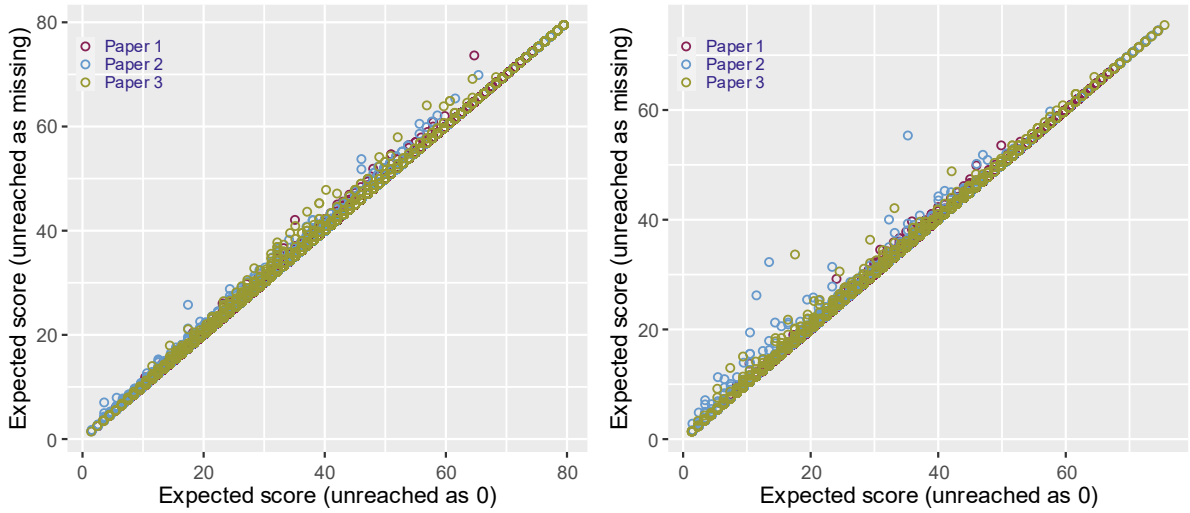

Figure 10 Relationship between expected paper score based on treating scores on unreached items as missing and expected paper score based on treating scores on unreached items as zeros for students taking the 3 higher tier maths papers (left) and the 3 foundation tier papers (right)

Figure 11 shows the percentage of students with differences between expected paper scores based on treating scores on unreached items as missing and the expected paper scores based on treating scores on unreached items as zeros (that is, $S_{n,exp} - S'_{n,exp}$) expressed as percentages of maximum paper marks (that is, relative differences) ≥1%, ≥2% and ≥3% of the maximum paper marks respectively in the 181 papers from the different subjects. Table 5 shows the corresponding values of range and mean for the different subjects.

For maths, the percentage of students with differences in expected scores greater than or equal to 1% of the maximum paper mark varied from 0.6% to 18.9% for the higher tier papers and from 1.1% to 10.9% for the foundation tier papers. Percentages of students with difference in expected scores greater than or equal to 2% and 3% of maximum paper marks for most of the higher tier papers were less than 3% and 2%, respectively. These were less than 5% and 3%, respectively, for most of the foundation tier papers.

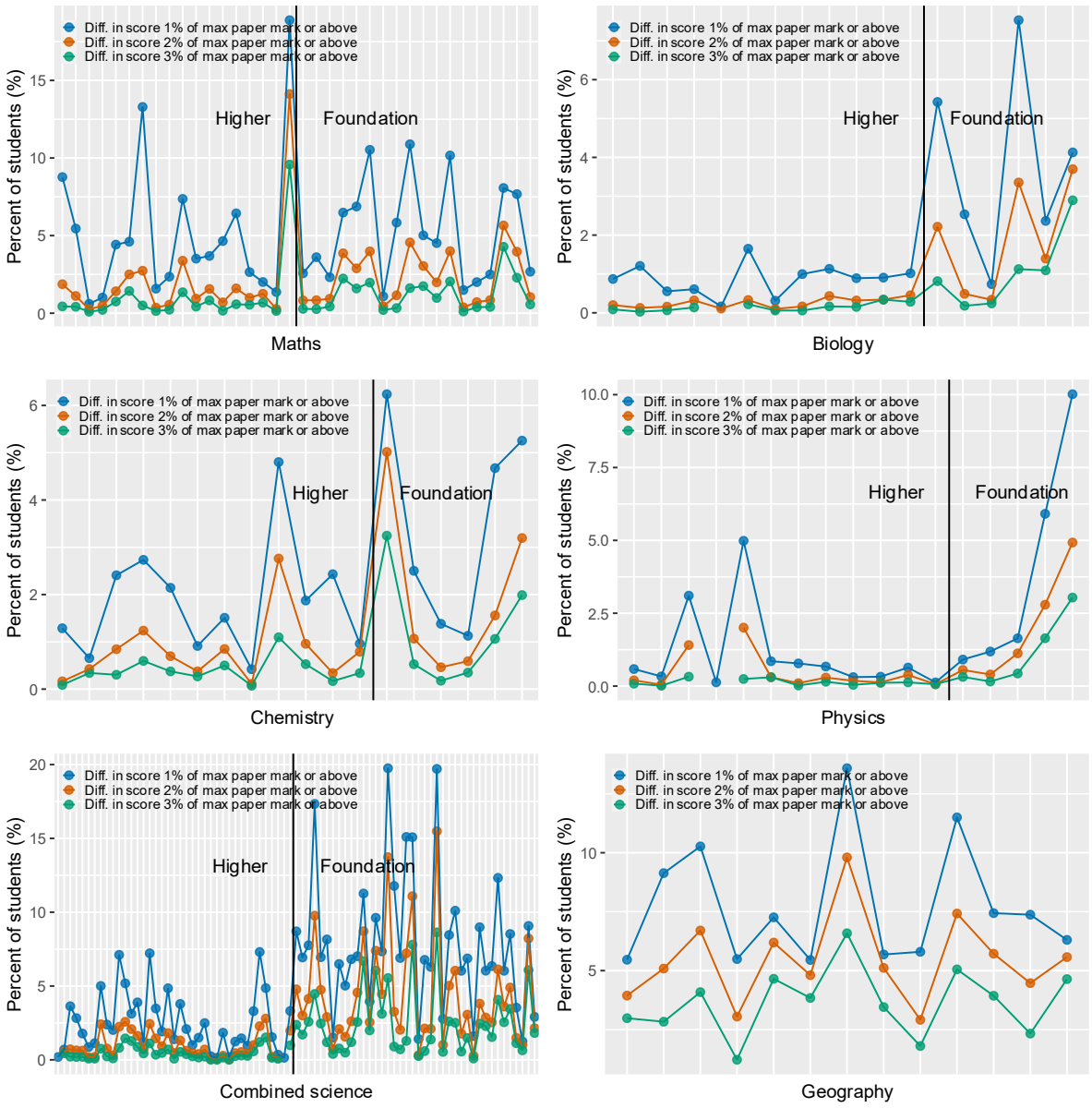

Figure 11 Distributions of percentage of students with difference in expected score ≥1%, ≥2% and ≥3% of the maximum paper score for the 181 papers analysed

Table 5 Range and mean of percentages of students with differences in expected scores greater than or equal to 1%, 2% and 3%, respectively, of the maximum paper marks for the papers from the different subjects

Subject	N	Range (mean) (exp. score diff ≥ 1%) (%)	Range (mean) (exp. score diff ≥ 2%) (%)	Range (mean) (exp. score diff ≥ 3%) (%)
Maths	38	0.60-18.88 (5.19)	0.25-14.12 (2.14)	0.09-9.58 (1.12)
Biology	18	0.16-7.53 (1.83)	0.10-3.70 (0.81)	0.00-2.90 (0.44)
Chemistry	18	0.42-6.24 (2.41)	0.11-5.02 (1.19)	0.07-3.25 (0.67)
Physics	17	0.13-10.02 (1.91)	0.00-4.93 (0.88)	0.00-3.04 (0.42)
Combined science	79	0.14-19.75 (5.27)	0.00-15.49 (2.68)	0.00-8.63 (1.47)
Geography	13	5.45-13.59 (7.75)	2.91-9.80 (5.45)	1.22-6.59 (3.65)

For the majority of the higher tier biology, chemistry, and physics papers, percentages of students with differences in expected scores greater than or equal to 1% of maximum paper marks were less than 2%. For the majority of the higher tier papers, percentages of students with differences in expected score greater than or equal to 2% of maximum paper marks were less than 1%. These were higher than 2% for the majority of the foundation tier papers. For most of the papers, percentages of students with differences in expected scores greater than or equal to 3% of maximum paper marks were less than 1%.

For combined science, for similar differences in expected scores, percentages of students for the higher tier papers were noticeably lower than those for the foundation tier papers. For most of the higher tier papers, percentages of students with differences in expected scores greater than or equal to 2% of maximum paper marks were less than 2%. These were higher than 4% for the majority of the foundation tier papers. For the majority the higher tier papers, percentages of students with differences in expected scores greater than or equal to 3% of maximum paper marks were less than 1%. These were above 3% for the majority of the foundation tier papers.

For the geography papers, percentages of students with differences in expected scores greater than or equal to 1%, 2% and 3% of maximum paper marks varied from 5.5%, 2.9% and 1.2% respectively to 13.6%, 9.8% and 6.6%. These values are generally higher than those for students with similar level of differences in expected scores for the majority of the papers from maths, biology, chemistry and physics.

Classification of students into different speeded groups

To investigate the magnitude of the impact of speededness on performance at the overall population level further, the distributions of differences in expected paper scores shown in Figure 11 were used to classify students into different speeded groups for each of the papers. The most-speeded group accounted for 10% of the highest values of the difference scores, the less-speeded group accounted for the next 15% of the highest values of the difference scores, and the least-speeded group accounted for 75% of the lowest values. Note that for some of the papers, percentages of students classified into the most- and less-speeded groups can be considerably smaller than 10% and 15% due to the limited number of discrete difference scores available. Further, for the majority of the biology, chemistry, physics and geography papers, there can only be 2 groups: the least- and most-speeded groups (the less-speeded group does not exist).

To see how the performance of the different speeded groups was affected by speededness, Figure 12 shows the means of differences between expected paper

scores on the 181 papers for the most- and less- speeded groups, expressed as percentages of maximum paper marks. (Note that the ranges of the y-axis in the graphs are different and that for some of the papers, the less-speeded group does not exist). The least-speeded groups were not included in the graphs as their performances were hardly affected. These mean difference values vary across the papers from the same subject and between the different subjects.

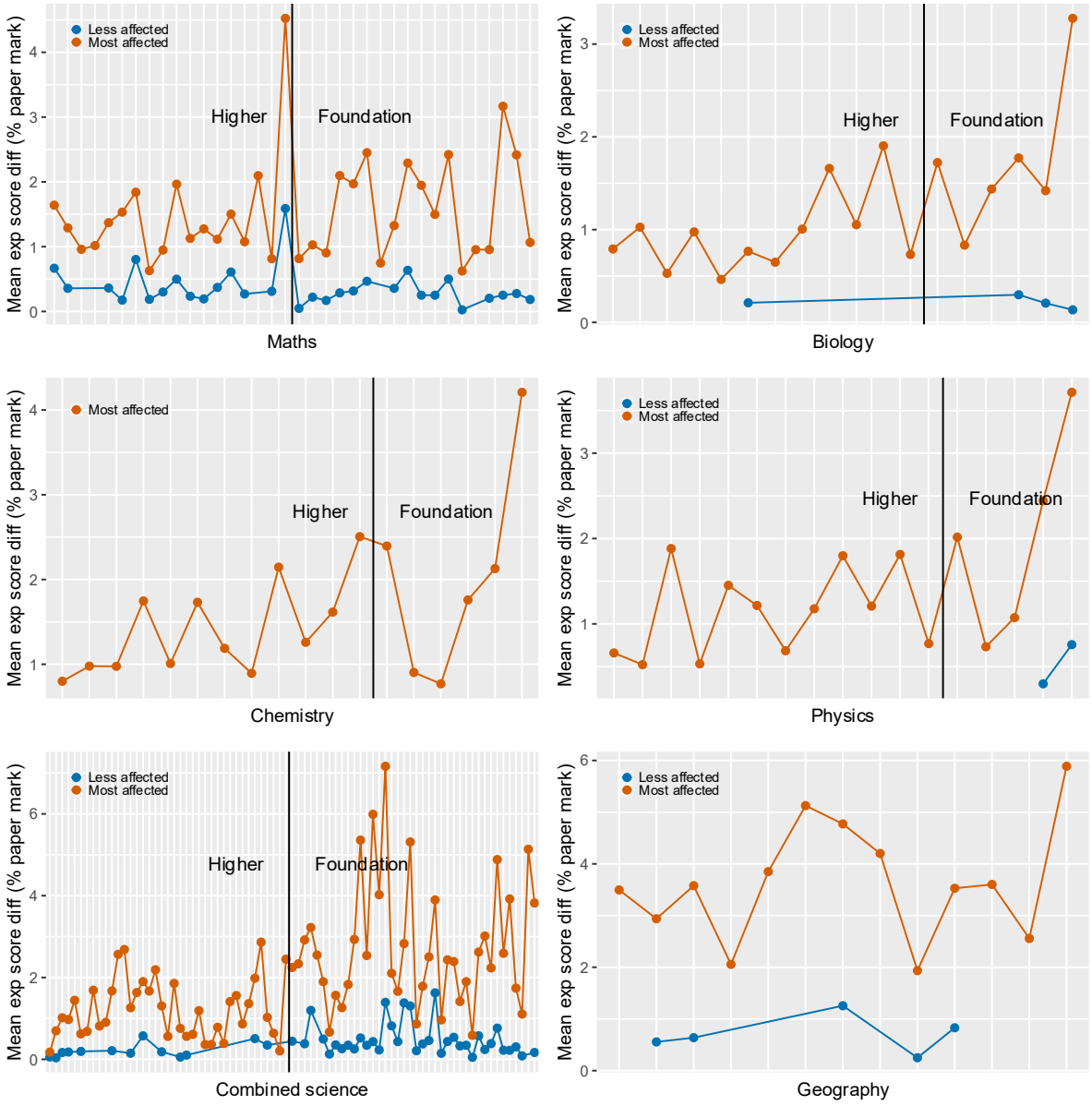

Figure 12 Distribution of means of differences in expected paper scores for the most- and less- speeded groups for the 181 papers analysed, expressed as percentages of maximum paper marks

For the most-speeded groups, those taking some of the geography and combined science foundation tier papers had the highest mean difference scores, with the values over 4% of the maximum paper marks. For those taking the other papers, most of them had mean difference scores less than 2% of the maximum paper marks. For most of the papers, the means of the mean differences in expected

scores for the less-speeded groups were less than 1% of the maximum paper marks. Table 6 lists the range and mean of the mean differences of the papers from the different subjects.

Table 6 Range and mean of mean differences in expected scores expressed as the percentage of maximum paper marks for the papers from the different subjects for the most-speeded groups

Subject	N	Range (%)	Mean (%)
Maths	38	0.62-4.52	1.54
Biology	18	0.46-3.28	1.22
Chemistry	18	0.77-4.21	1.61
Physics	17	0.52-3.71	1.39
Combined science	79	0.17-7.16	2.00
Geography	13	1.93-5.89	3.66

Impact of speededness across ability range

To further see how students with different abilities are affected by speededness, Figure 13 shows the distributions of scores on the 3 higher tier maths papers and the 3 foundation tier papers for the different speeded groups. Note that the scores on papers will have been affected by speededness. For the 6 papers, students from the most-speeded groups have higher average paper scores than students from the less-speeded groups (note that foundation Paper 2 only has 2 groups: the least- and most-speeded groups). Students from the least-speeded groups have the highest average paper scores. The distributions in Figure 13 are consistent with the relationships observed in Figure 10 which compares expected paper scores based on treating scores on unreached items as missing and those based on treating scores on unreached items as zeros.

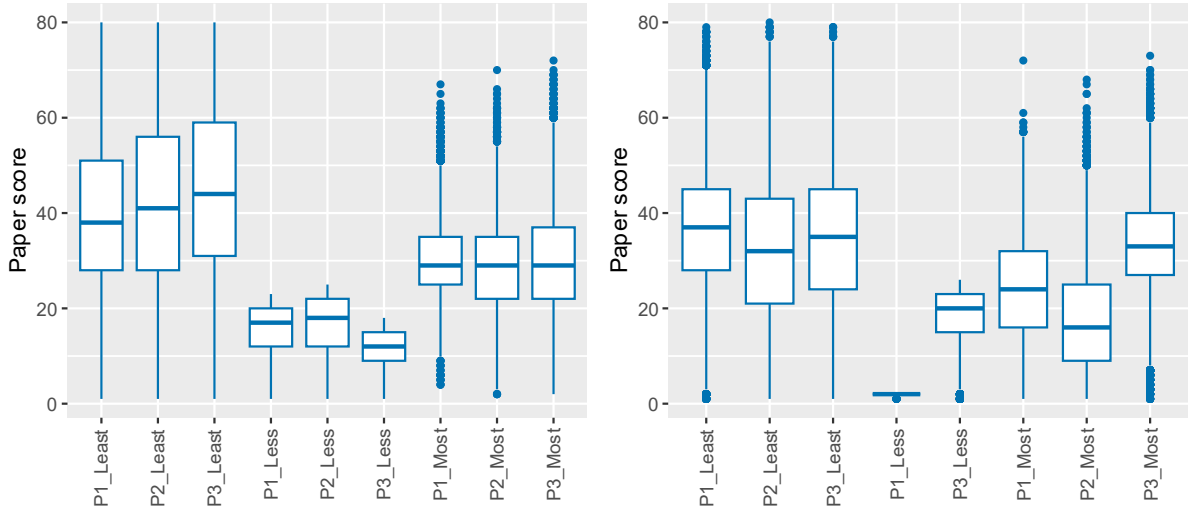

Figure 13 Distributions of scores of students from different speeded groups taking the 3 higher tier maths papers (left) and 3 foundation tier papers (right)

To see more clearly how students with different abilities are affected by speededness, Figure 14 shows the distribution of mean Rasch ability based on treating scores on unreached items as missing for the different speeded groups across the 181 papers. It is clear from Figure 14 that, for almost all the papers analysed, the mean abilities of the least-speeded groups are higher than those of the most-speeded groups which in turn are higher than those of the less-speeded groups. This suggests that, on average, the most and least able students are less likely to be affected by speededness than students with average abilities.

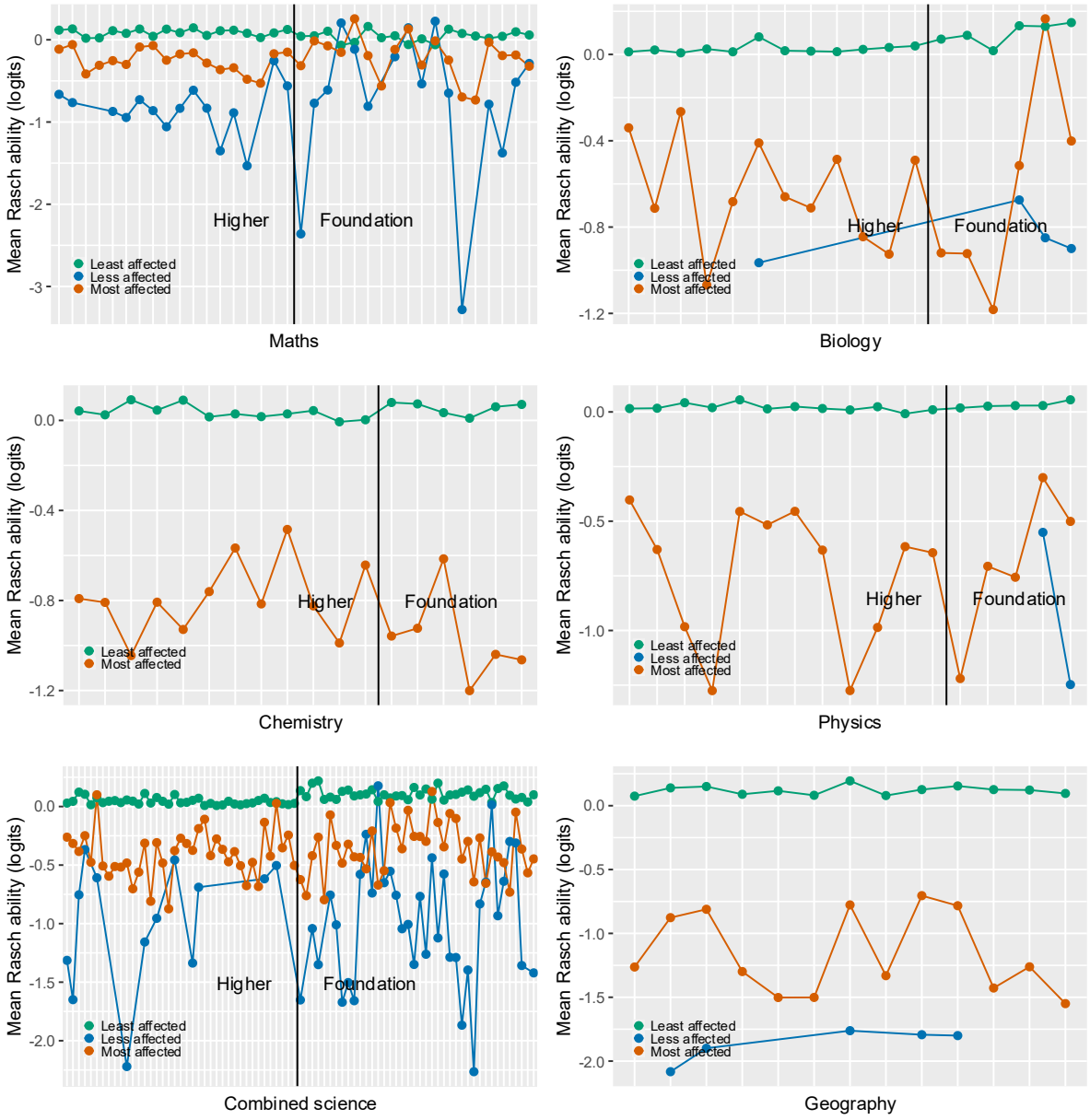

Figure 14 Distributions of mean Rasch abilities based on treating scores on unreached items as missing for the different speeded groups across the 181 papers analysed

Consistency of classification between papers in the same specification

To see how consistently students were classified between papers in the same specification, Table 7 lists proportions of students who were classified into the most-speeded group by pairs of papers from the same specification. For the maths papers, consistency in classification for the most-speeded groups varied from 0.20 to 0.28, with a mean of 0.27. For biology, chemistry, physics and combined science papers, this varied from 0.20 to 0.47, 0.15 to 0.44, 0.06 to 0.44, and 0.18 to 0.17 respectively. The consistency for the geography papers varied from 0.38 to 0.63, with a mean of 0.50. These values are noticeably higher than those for the papers from the other 5 subjects.

Table 7 Consistency of classifying students into the most-speeded groups by pairs of papers from the different subjects

Subject	Specification	Tier	Papers 1&2 (3&4 or 5&6)	Papers 1&3	Papers 2&3
Maths	SP1	H	0.25-0.30	0.26-0.28	0.20-0.25
	SP1	F	0.24-0.38	0.22-0.24	0.26-0.28
	SP2	H	0.27-0.29	0.26-0.29	0.22-0.32
	SP2	F	0.27-0.30	0.26-0.33	0.27-0.30
Biology	SP1	H	0.20,0.21		
	SP1	F	0.26,0.40		
	SP2	H	0.22,0.27		
	SP2	F	0.26		
	SP3	H	0,46,0.47		
Chemistry	SP1	H	0.23		
	SP1	F	0.26,0.40		
	SP2	H	0.34,0.36		
	SP2	F	0.40		
	SP3	H	0.14-0.44		
Physics	SP1	H	0.15,0.28		
	SP1	F	0.26		
	SP2	H	0.23,0.44		
	SP2	F	0.31		
	SP3	H	0.06,0.39		
Combined science	SP1	H	0.22-0.39		
	SP1	F	0.30-0.43		
	SP2	H	0.18-0.34		
	SP2	F	0.23-0.31		
	SP3	H	0.35-0.47		
	SP3	F	0.28-0.34		
	SP4	H	0.24-0.27		
	SP4	F	0.31-0.42		
Geography	SP2		0.44,0.46	0.40, 0.44	0.38, 0.41
	SP3		0.52,0.57	0.51, 0.62	0.60, 0.63

Except for the geography papers, the level of consistency in classification between the papers shown in Table 7 is relatively low. This may partly reflect the fact that

different papers in a specification are designed to assess different constructs and students' knowledge and skills vary across the papers. As a result, students may be affected differently by speededness when taking different papers. Further, the classification of students taking a paper into different speededness groups involves the estimation of Rasch abilities and expected scores and subsequent classification of students into different speeded groups of arbitrary size, which will have error built in and reduce the consistency in the classification of students between papers. Another factor that could have affected the level of consistency in the classification is the various extra time adjustments received by some of the students when taking the exams. For most of the papers and students analysed, it seems that speededness could be thought of more as a property of the paper than a property of the student.

3.5 Relationship between speededness and students' background characteristics

We have seen earlier in this report that item omission rate varies between different population subgroups. This section explores further the relationship between speededness and students' background characteristics using multiple multinomial logistic regression analysis.

Table B1 in Appendix B show values of the multinomial logistic regression model coefficients between different speeded groups classified based on the partial credit Rasch model for the various variables which are significantly different from zero (at $p<0.05$ level) for the 181 papers from the different subjects (the least-speeded group was used as the reference group). Note that it is not appropriate to compare the relative importance of the different independent variables directly as the scales of the variables can be different.

As an example, Figure 15 compares the model coefficients and classification probabilities against ability for the 3 higher tier maths papers and the 3 foundation tier papers discussed previously. The graphs in the first and third columns in Figures 16 compare model coefficients between speeded groups for the 6 papers. The 95% confidence intervals of the parameter estimates are also shown in the graphs. These graphs can be used to assess whether the effects associated with individual variables are significantly different between the different speeded groups.

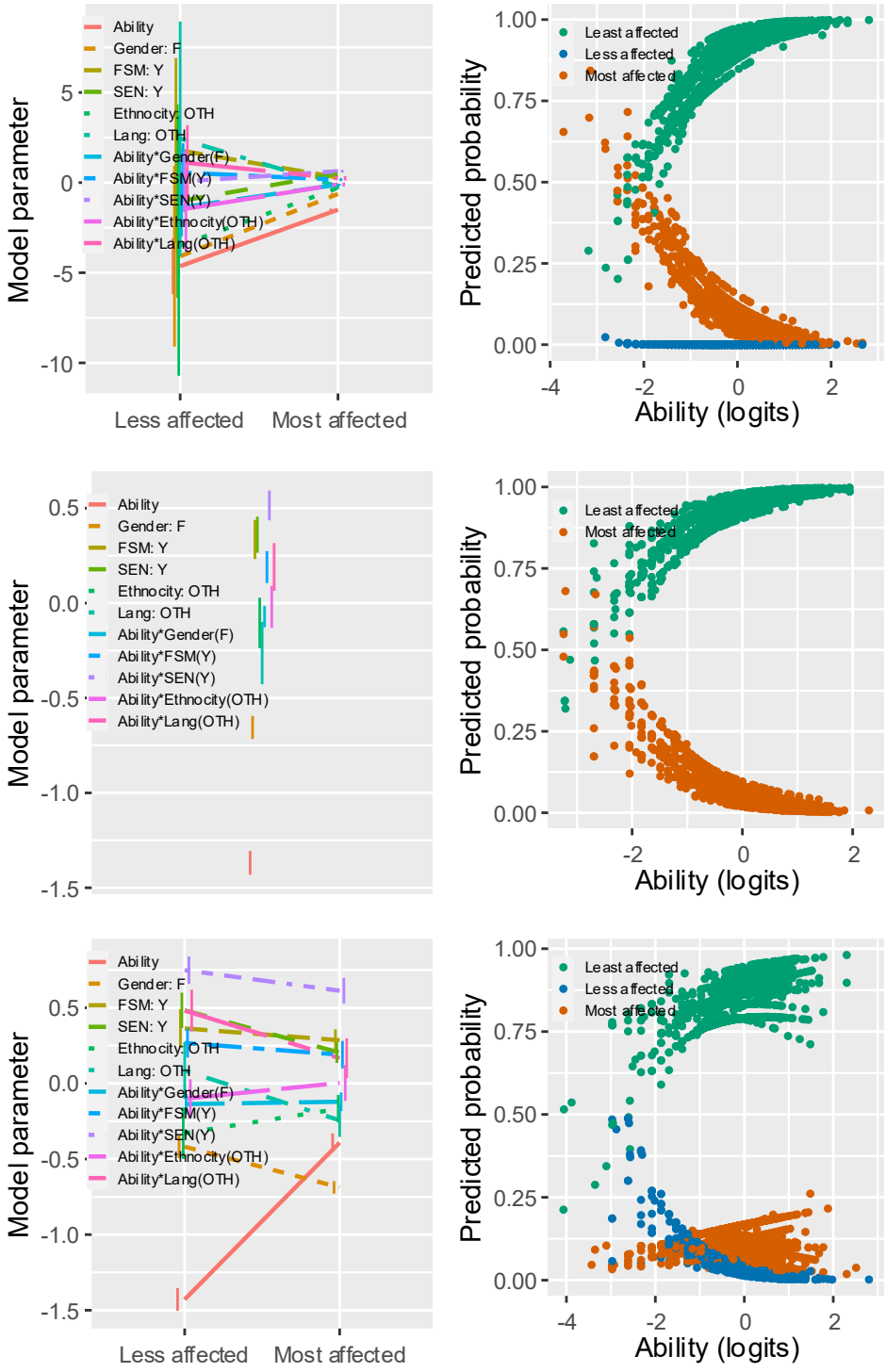

Figure 15 Model parameters (column 1) from multinomial logistic regression and distributions of probabilities of classification against ability (column 2) for the 3 foundation tier maths papers

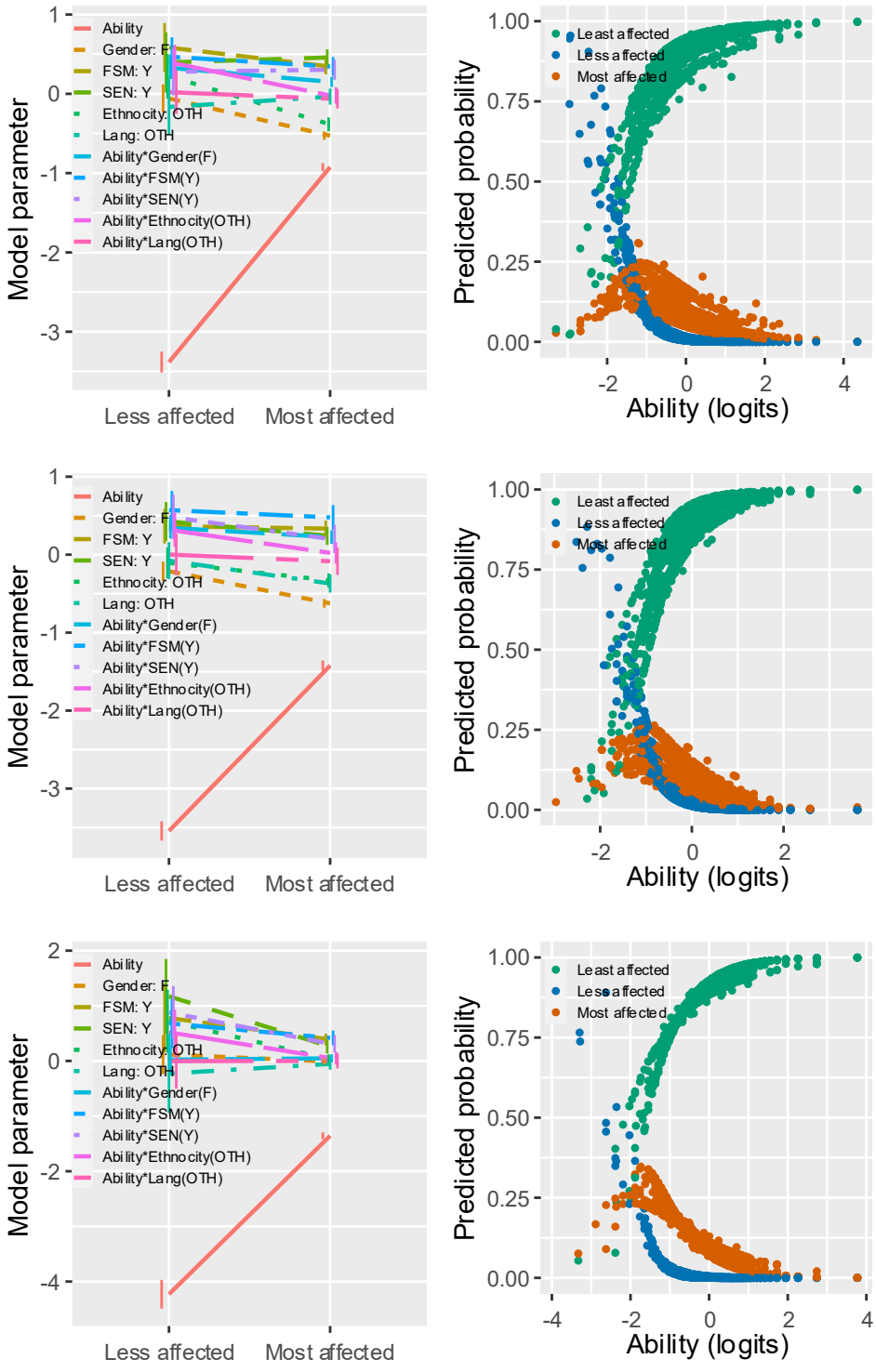

Figure 16 Model parameters (column 1) from multinomial logistic regression and distributions of probabilities of classification against ability (column 2) for the 3 higher tier maths papers

The graphs in the second column of Figures 15 and 16 show how the probabilities of individual students being classified into different speeded groups vary with their ability. For the same speeded group, there are several sets of probability curves which reflect different values of the model coefficients for the other variables,

including sex, eligibility for FSMs, SEN status, ethnicity and others. For all the 6 papers, the probability of being classified into the least-speeded group increases with ability, whereas that of being classified into the most-speeded group decreases (also see earlier discussion). For the 3 higher tier papers and one of the foundation tier papers (paper P3), the probabilities of being classified into the most-speeded group increase with ability at the lower ability end and then decrease towards the higher ability end. For the other 2 foundation tier papers, the probabilities of being classified into the most-speeded group decrease with ability monotonically. For the rest of the papers analysed, similar probability distributions are observed.

For the papers listed in Table B1, the model parameters associated with the ability variable are negative and significantly different from zero for all the papers. Other things being equal, the probabilities of being classified into the least-speeded groups increase with ability, whereas those of being classified into the most-speeded and less-speeded groups decrease. For the other categorical variables, if the coefficient is positive, the probability of being classified into the less-speeded and most-speeded groups for the focal group is higher than for the reference group (other things being equal). If, on the other hand, the coefficient is negative, the probability of being classified into the less-speeded and most-speeded groups for the focal group is lower than for the reference group.

For the majority of the papers, values of the model parameter for sex are negative and significantly different from zero. Other things being equal, boys (the reference group) generally have higher probabilities of being classified into the most- and less-speeded groups than girls (the focal group). Model parameters for eligibility for FSM and SEN status are positive and significantly different from zero for some of the papers, with those eligible for FSM and with SEN having higher probabilities of being classified into the most-speeded groups than those not eligible for FSM and those without SEN (other things being equal). For language and ethnicity, model parameters for the majority of the papers are not significantly different from zero, suggesting similar impact from speededness with respect to these variables. These findings are broadly consistent with the findings presented earlier.

For some of the maths, combined science and geography papers, model coefficients for the interaction terms between ability and sex, eligibility for FSM and SEN status are also statistically significantly different from zero. For FSM and SEN, the signs of the interaction coefficients are consistent with those of the group variables. For sex, some of the signs are consistent with those of the sex variable, but some are different.

The inclusion of students who had received extra time adjustments when taking the papers in the analysis is likely to have affected the relationship between classification and students' background characteristics.

4. Discussion

In this study, the distributions of item omission rate for 181 GCSE exam papers from mathematics, biology, chemistry, physics, combined science, and geography were estimated. It was assumed that omissions of end-of-test items in these papers were caused by students not having sufficient time to answer all the questions in the papers.

The omission rates of items near the end of the papers varied across the papers from the 6 different subjects. The omission rates for the biology, chemistry and physics papers were generally lower than those for the maths, combined science and geography papers. For most of the maths and combined science papers, percentages of items with omission rate greater than or equal to 2% were less than 10%. For most of the biology, chemistry and physics papers, these were less than 5%. For the majority of the geography papers, these were greater than 15%. For the 5 tiered subjects, higher tier papers generally had lower item omission rates than the foundation tier papers. The maximum item omission rates (the omission rates of the last items in the papers) varied from 3.3 to 51.5% (mean 20.8%) for mathematics, 1.8 to 25.1% (mean 7.4%) for biology, 1.0 to 8.4% (mean 4.3%) for chemistry, 0.5 to 15.8% (mean 4.7%) for physics, 2.5 to 58.1% (mean 14.5%) for combined science, and 5.8 to 15.4% (mean 9.7%) for geography. There is substantial variation in the apparent degree of speededness across the papers from the different subjects.

There was considerable variability in item omission rate between the different population groups in the papers from the different subjects. Students with SEN and those eligible for FSMs had the highest omission rates. Female students generally had lower item omission rates than male students.

For each paper studied, the partial credit Rasch model was used to study the impact of possible speededness on students' performance and the results were used to classify students into 3 groups: the least-speeded group, the less-speeded group and the most-speeded group, accounting for 75%, 15% and 10% respectively of the total number of students taking the paper. For the least-speeded groups, the mean of differences between the expected paper scores based on treating scores on unreached items as missing and the expected paper scores based on treating scores on unreached items as zeros was close to zero for all the 181 papers analysed. For the less-speeded groups, for most of the papers, the mean of differences in expected paper scores was less than 0.5% of maximum paper mark. For the most-speeded groups, the mean differences in expected scores ranged from about 0.62 to 4.52% (mean 1.54%) of maximum paper marks for mathematics, 0.46 to 3.28% (mean 1.22%) for biology, 0.77 to 4.21% (mean 1.61%) for chemistry, 0.52 to 3.71% (mean 1.39%) for physics, 0.17 to 7.16% (mean 2.00%) for combined science, and 1.93 to 5.89% (mean 3.66%) for geography. The most and least able

students were less likely to be affected by speededness than students with average abilities.

The consistency of classifying students into the most-speeded group by pairs of papers varied from 0.20 to 0.38 for maths, 0.20 to 0.47 for biology, 0.15 to 0.44 for chemistry, 0.06 to 0.44 for physics, 0.18 to 0.47 for combined science, and 0.38 to 0.60 for geography. Except for the geography papers, the relatively low consistency between the papers from the other subjects likely reflects differences in constructs assessed by the different papers in a specification, errors involved in producing the classifications, and other factors. This suggests that, for the papers and students analysed, speededness could be thought of more as a property of the paper than a property of the student.

Analysis using multinomial logistic regression indicated that the model parameters associated with the ability variable were significantly different from zero for almost all the papers analysed. Other things being equal, the probabilities of students being classified into the least-speeded groups increased with increasing ability, whereas those of being classified into the most-speeded groups decreased. For most of the papers, values of model parameter for sex were significantly different from zero. Other things being equal, male students seemed to have higher probabilities of being classified into the most- and less-speeded groups than female students. Model parameters for eligibility for FSM and SEN status were significantly different from zero for some of the papers, with those eligible for FSM and with SEN having higher probabilities of being classified into the most-speeded groups than those not eligible for FSM and those without SEN (other things being equal). For language and ethnicity, model parameters for the majority of the papers were not significantly different from zero, suggesting similar impact from speededness with respect to these variables. For some of the maths, combined science and geography papers, model coefficients for the interaction terms between ability and sex, eligibility for FSM and SEN status were also statistically significantly different from zero.

There are, however, important limitations in this study. The method used to estimate item omission rate distribution and to identify items affected by speededness and the use of the partial credit Rasch model to classify students into different speededness groups involved a number of assumptions which might not have been fully met by the datasets analysed. This could have over-estimated (or under-estimated) the extent of speededness in the papers and introduced errors in classification. This in turn could have reduced consistency in classifications between papers from the same specification. The observed response patterns for the end-of-test items in the various papers could have resulted from a range of factors other than speededness, such as changes in topic areas, the perceived level of difficulty, motivation, effort, and fatigue. Further, some students may not have answered questions in the order of sequence in the papers. The assumption of least effect of speededness for students who have valid marks for all items in a paper might also not hold.

Another limitation is that the data analysed contained students who had received extra time when taking these exams, and these students could not be identified in the data. This would have affected the estimation of the effect from possible speededness in these exams (the effect was likely to have been underestimated), consistency of classification between papers, and the relationship between classification and students' characteristics.

Finally, this approach is limited to exam papers that have a reasonable number of questions requiring short to moderate length answers. Subjects in which responses take the form of essays are not amenable to this kind of analysis.

Further analysis of changes in the nature of the first few items assumed to be affected by speededness in the papers (for example, their level of demand or difficulty or their topic areas) could provide useful insight into the possible reasons for the observed response patterns. Post-exam questionnaires of students' experience in answering questions (for example motivation, effort, question answering strategy and tactics, time pressure, and so on) could provide useful information to help explain unusual responses associated with end-of-test items in these examinations. Analysis restricted to those students who had not received extra time would also allow more accurate estimation of the effect from potential speededness in the papers investigated here.

Notwithstanding the limitations of the methods used in this study, there appears to have been a degree of speededness in the exam papers investigated, and students with different abilities were affected differently. Overall, the degree of speededness in the papers analysed does not seem to be high, with most students affected only losing a very small number of marks due to possible time pressure. Yet it should be noted that some papers were more speeded than others, and for a small number of students the effect of time limits on their performance may have been substantial. Speed of responding is not part of the stated construct for GCSEs and A levels, and high levels of speededness could compromise their validity and fairness. This form of analysis can help inform decisions regarding whether the number and demand of items in question papers and their associated time limits are appropriate to avoid speed of working becoming a significant factor affecting student performance. In addition, speededness may vary year by year across papers within a single qualification, and so an analysis such as this may lend itself to the ongoing monitoring of question paper functioning.

References

AERA (American Educational Research Association), APA (American Psychological Association), and NCME (National Council on Measurement in Education), eds. (2014). *Standards for Educational and Psychological Testing*. Lanham, MD: American Educational Research Association.

Andrich, D. (2015). The problem with the step metaphor for polytomous models for ordinal assessments. *Educational Measurement: Issues and Practice* 34, 8–14.

Bolt, D. M., Cohen, A. S., & Wollack, J. A. (2002). Item parameter estimation under conditions of test speededness: Application of a mixture Rasch model with ordinal constraints. *Journal of Educational Measurement* 39, 331–348.

Borter, N., Schlegel, K. and Troche, S. (2023). How Speededness of a reasoning test and the complexity of mental speed tasks influence the relation between mental speed and reasoning ability. *Journal of Intelligence* 11: 89. https://doi.org/10.3390/jintelligence11050089

Cao, J. and Stokes, S. (2008). Bayesian IRT guessing models for partial guessing behaviors. *Psychometrika* 73, 209–230.

Cho, S. and Cohen, A. (2010). A multilevel mixture IRT model with an application to DIF. *Journal of Educational and Behavioral Statistics* 35, 336–370.

Cintron, D. (2021). Methods for measuring speededness: chronology, classification, and ensuing research and development. ETS report ETS RR–21-22, ETS. Available online at: Methods for Measuring Speededness: Chronology, Classification, and Ensuing Research and Development - Cintron - 2021 - ETS Research Report Series - Wiley Online Library

Evans, F. and Reilly, R. (1972). A study of speededness as a source of test bias. *Journal of Educational Measurement* 9, 123-131.

Goegebeur, Y., De Boeck, P., Molenberghs, G. anddel Pino, G. (2006). A local-influence based diagnostic approach to a speeded item response theory model. *Journal of the Royal Statistical Society. Series C (Applied Statistics)* 55, 647–676.

Goegebeur, Y., De Boeck, P., Wollack, J., and Cohen, A. (2008). A speeded item response model with gradual process change. *Psychometrika 73*, 65–87.

Holmes, S. (2025). *Time limits and speed of working in assessments: When, and to what extent, should speed of working be part of what is assessed?* Ofqual research report 25/7267/2. Coventry, UK: Ofqual. Retrieved from https://www.gov.uk/government/publications/time-limits-and-speed-of-working-in-assessments

Hong, M. and Cheng, Y. (2019). Clarifying the Effect of Test Speededness. *Applied Psychological Measurement* 43, 611–623.

JCQ (2020). Adjustments for candidates with disabilities and learning difficulties – access arrangements and reasonable adjustments. Available online at: https://www.jcq.org.uk/wp-content/uploads/2020/08/AA-regs-2020-2021-version-for-website.pdf

JCQ (nd). Access arrangements, reasonable adjustments and special consideration. Available online at: https://www.jcq.org.uk/exams-office/access-arrangements-and-special-consideration/

Jin, K. and Wang, W. (2014). Item Response Theory Models for Performance Decline During Testing. *Journal of Educational Measurement* 51, 178-200.

Lu, Y. and Sireci, S. (2007). Validity Issues in Test Speededness. *Educational Measurement Issues and Practice* 26, 29-37.

Lunn, D., Thomas, A., Best, N., and Spiegelhalter, D. (2000). WinBUGS – a Bayesian modelling framework: concepts, structure, and extensibility. *Statistics and Computing* 10, 325–337.

Masters, G. (1982). A Rasch model for partial credit scoring. *Psychometrika* 47, 149–174.

Mroch, A., Bolt, D. and Wollack, J. (2005, April). *A new multi-class mixture Rasch model for test speededness*. Paper presented at the annual meeting of the National Council on Measurement in Education, Montreal, Canada.

Ripley, B. and Venables, W. (2021). Package 'nnet'. Available online at: https://cran.r-project.org/web/packages/nnet/nnet.pdf

Robitzsch, A., Kiefer, T. and Wu, M. (2022). Package 'TAM'. Available online at: https://cran.r-project.org/web/packages/TAM/TAM.pdf

Sen, S. and Cohen, A. (2019). Applications of mixture IRT Models: a literature review. *Measurement: Interdisciplinary Research and* Perspectives 17, 177-191.

Sturtz, S. and Ligges, U. (2015). Package 'R2WinBUGS'. Available online at: https://cran.r-project.org/web/packages/R2WinBUGS/R2WinBUGS.pdf

Suh, Y., Cho, S. and Wollack, J. (2012). A comparison of item calibration procedures in the presence of test speededness. *Journal of Educational Measurement* 49, 285–311.

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.

Wheadon, C. (2011). An item response theory approach to the maintenance of standards in public examinations in England, Durham theses, Durham University. Available at Durham E-Theses Online: http://etheses.dur.ac.uk/615/

Wise, S. L. (1996, April). *A persistence model of motivation and test performance*. Paper presented at the annual meeting of the American Educational Research Association, New York, NY.

Wright, B. and Masters, G. (1982). *Rating scale analysis: Rasch measurement*. Chicago: MESA Press.

Yamamoto, K. and Everson, H. (1997). Modeling the effects of test length and test time on parameter estimation using the HYBRID model. In J. Rost and R. Langeheine (Eds.), *Applications of latent trait and latent class models in the social sciences* (pp. 89–98). Munster, Germany: Waxmann.

Yu X. and Cheng, Y. (2022). A comprehensive review and comparison of CUSUM and Change-Point-Analysis methods to detect test speededness. *Multivariate Behav Res*. 57,112-133.

Appendix A: Average percentage marks lost from unattempted items

Walland (2024) investigated the level of speededness in 340 GCSE written papers administered from 2009 to 2016 using average percentage marks lost from unreached items at the end of the paper of the students (that is, average maximum available marks associated with unreached items expressed as percentage of the maximum available paper mark). In her study, students taking a paper were divided into 4 ability groups (with similar number of students in each group). She suggested that in a speeded paper where students experience time pressure, there would be strings of items which are unattempted near the end of the paper for higher achieving students (or noticeable level of average percentage marks lost from unreached items). This is because, for lower ability students, unattempted items may result from the items being too difficult and it is only the more able students who would attempt every question they reached and provide a relatively uncontaminated estimate of speededness. For most of the papers she analysed, the average percentage marks lost associated with unreached items for the top 25% high performing students was small. She therefore suggested that most of the exam papers she analysed were unlikely to be speeded.

Similar analysis was conducted for the 181 GCSE exam papers investigated in the present study. Students taking an exam paper are divided into 4 ability groups, Q1, Q2, Q3, and Q4, with Q1 representing 25% of the lowest performing students and Q4 25% of the highest performing students, based on their overall paper scores. Figure A1 below shows the distributions of average percentage marks associated with unreached items for each of the ability groups taking the papers from the 6 subjects. For tiered subjects, the higher and foundation tiers are reported separately. Table A1 lists the range and mean of the average percentage marks lost for individual ability groups for individual subjects and the 5 tiered subjects combined. The patterns seen in Figure A1 and Table A1 for the papers analysed here are broadly similar to those for the papers investigated by Walland (2024). As is clear from Figure A1 and Table A1, average percentage marks associated with unreached items decrease with increasing ability. Average percentage marks lost for students in group Q4 (the highest performing group) are much smaller than those for students in group Q1 (the lowest performing group). For the tiered subjects, foundation tier students have higher average percentage marks lost than higher tier students for every ability group. Except for some of the foundation tier papers in maths and combined science, average percentage marks lost from unreached items for the top 25% high performing students is well below 1% of the maximum paper marks. For some of the foundation tier maths and combined science papers, the average

percentage marks associated with unreached items for the top 25% highest performing students are close to or higher than 1%. Based on average percentage marks lost from unreached items for high performing students, it would appear that the level of speededness is low for most of the papers investigated here. For the 5 tiered subjects, overall, marks lost from unreached items shown in Table A1 are slightly higher than those reported by Walland. For the higher tier papers these are close to or very slightly lower than those she reported. For the foundation tier papers, they are higher.

However, a test that is not speeded for high ability students may be speeded for moderate ability students, depending on its overall difficulty relative to the ability of the students taking the test. GCSE exam papers are designed to cover a wide range of abilities. Given that students in Q3 and Q2 account for 50% of those taking the paper, it may be more appropriate to use their response patterns to evaluate if the paper is speeded. Figure A1 indicates noticeable level of average percentage marks lost from unreached items for groups Q3 and Q2 for the majority of the papers analysed, suggesting that there is a degree of speededness in these papers.

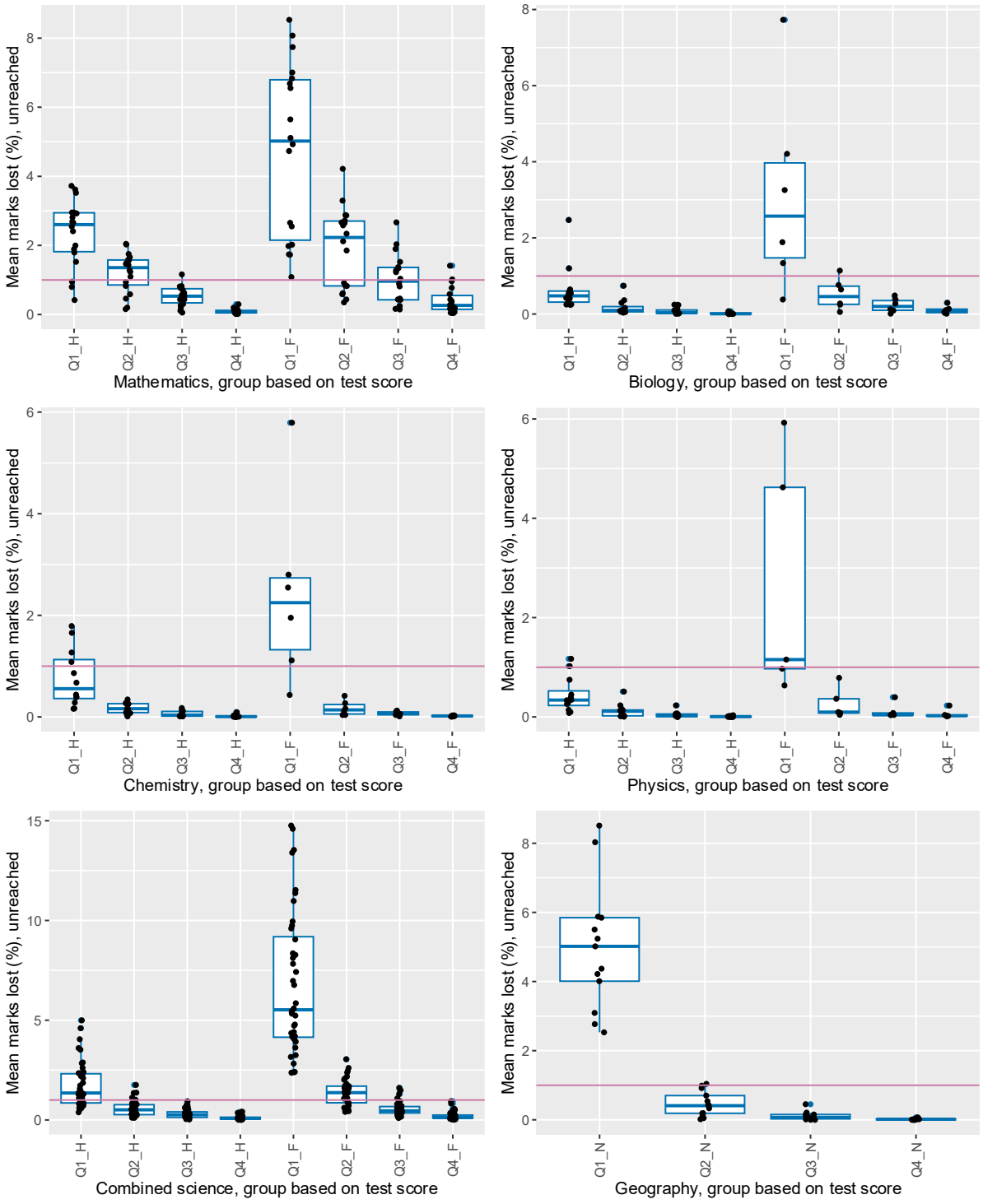

Figure A1 Average percentage marks lost from **unreached items** for each of the ability groups taking the papers from the 6 subjects investigated. A horizontal line at y=1.0 is drawn in the graphs

Table A1 Mean and standard deviation of average percentage marks lost from unreached items and all unattempted items for students in each ability group for individual subjects studied

Subject	Tier	N	Group	Unreached items			All unattempted items		
				Range	Mean	SD	Range	Mean	SD
Maths	H	18	Q1	0.42-3.72	2.34	0.96	8.36-18.96	11.77	2.89
		18	Q2	0.16-2.05	1.21	0.58	4.14-10.42	6.21	1.82
		18	Q3	0.06-1.16	0.53	0.28	1.51-5.09	2.86	1.02
		18	**Q4**	**0.01-0.30**	**0.09**	**0.07**	**0.17-0.87**	**0.50**	**0.21**
	F	18	Q1	1.08-8.53	4.76	2.52	15.5-31.27	21.83	4.67
		18	Q2	0.35-4.22	1.93	1.16	5.7-17.77	10.48	3.41
		18	Q3	0.15-2.67	1.01	0.72	2.68-11.09	5.73	2.36
		18	**Q4**	**0.04-1.41**	**0.41**	**0.39**	**1.08-6.12**	**2.47**	**1.38**
Biology	H	12	Q1	0.24-2.47	0.66	0.63	2.53-8.92	3.85	1.71
		12	Q2	0.04-0.74	0.18	0.21	0.56-3.66	1.18	0.84
		12	Q3	0.01-0.25	0.08	0.09	0.14-1.61	0.50	0.41
		12	**Q4**	**0-0.07**	**0.02**	**0.02**	**0.03-0.61**	**0.15**	**0.16**
	F	6	Q1	0.38-7.73	3.13	2.63	9.03-24.68	16.04	5.62
		6	Q2	0.05-1.14	0.52	0.40	1.42-6.12	4.11	1.86
		6	Q3	0.01-0.49	0.23	0.18	0.53-3.15	1.90	0.88
		6	**Q4**	**0.01-0.3**	**0.10**	**0.10**	**0.24-1.6**	**0.82**	**0.44**
Chemistry	H	12	Q1	0.16-1.79	0.76	0.57	2.95-10.52	5.77	2.09
		12	Q2	0.01-0.34	0.17	0.11	0.90-2.62	1.60	0.51
		12	Q3	0.01-0.17	0.06	0.06	0.23-0.82	0.56	0.18
		12	**Q4**	**0-0.09**	**0.02**	**0.03**	**0.03-0.26**	**0.14**	**0.07**
	F	6	Q1	0.43-5.79	2.44	1.87	11.45-22.61	15.11	4.06
		6	Q2	0.04-0.42	0.17	0.15	3.11-6.69	4.38	1.32
		6	Q3	0.01-0.12	0.07	0.04	1.61-4.50	2.55	1.08
		6	**Q4**	**0.01-0.03**	**0.02**	**0.01**	**0.63-2.71**	**1.18**	**0.77**
Physics	H	12	Q1	0.08-1.17	0.44	0.35	2.43-8.42	4.56	1.99
		12	Q2	0.01-0.51	0.13	0.14	0.46-2.54	1.27	0.61
		12	Q3	0.00-0.23	0.05	0.06	0.11-1.56	0.55	0.42
		12	**Q4**	**0.00-0.03**	**0.01**	**0.01**	**0.03-0.35**	**0.14**	**0.10**
	F	5	Q1	0.63-5.93	2.66	2.44	9.1-19.99	13.64	3.97
		5	Q2	0.04-0.79	0.27	0.31	1.51-6.15	3.35	1.77
		5	Q3	0.04-0.40	0.12	0.16	0.64-2.84	1.70	0.79
		5	**Q4**	**0.01-0.23**	**0.06**	**0.09**	**0.29-1.28**	**0.76**	**0.38**
Combined science	H	39	Q1	0.38-4.99	1.76	1.17	3.15-21.64	9.36	4.37
		39	Q2	0.11-1.76	0.56	0.38	1.14-9.82	3.94	1.87
		39	Q3	0.03-0.93	0.30	0.22	0.61-4.90	2.22	1.05
		39	**Q4**	**0.01-0.42**	**0.10**	**0.09**	**0.2-2.18**	**0.85**	**0.52**
	F	40	Q1	2.37-14.76	6.85	3.55	10.12-35.83	22.34	6.95
		40	Q2	0.40-3.04	1.37	0.64	2.29-15.6	7.74	3.24
		40	Q3	0.10-1.61	0.56	0.35	0.93-9.42	4.03	1.96
		40	**Q4**	**0.01-0.94**	**0.21**	**0.21**	**0.34-4.87**	**1.80**	**1.04**
Geog		13	Q1	2.53-8.51	5.00	1.83	13.73-24.75	18.85	3.12
		13	Q2	0.02-1.04	0.45	0.36	0.68-3.74	2.60	1.06
		13	Q3	0.00-0.45	0.11	0.12	0.18-1.52	0.83	0.41
		13	**Q4**	**0.00-0.07**	**0.02**	**0.02**	**0.05-0.57**	**0.26**	**0.16**
M, B, C, P, CS	H	93	Q1	0.08-4.99	1.43	1.14	2.43-21.64	8.03	4.33
		93	Q2	0.01-2.05	0.53	0.53	0.46-10.42	3.38	2.35
		93	Q3	0-1.16	0.25	0.26	0.11-5.09	1.69	1.26
		93	Q4	0-0.42	0.07	0.08	0.03-2.18	0.51	0.48
		93	**Overall**	**0.03-1.94**	**0.59**	**0.49**	**0.78-9.83**	**3.50**	**2.07**
	F	75	Q1	0.38-14.76	5.42	3.46	9.03-35.83	20.56	6.63
		75	Q2	0.04-4.22	1.27	0.93	1.42-17.77	7.55	3.72

		75	Q3	0.01-2.67	0.57	0.53	0.53-11.09	3.99	2.26
		75	Q4	0.01-1.41	0.23	0.27	0.24-6.12	1.76	1.16
		75	**Overall**	**0.12-5.18**	**1.94**	**1.19**	**2.89-16.34**	**8.71**	**3.31**
	Combined	168	Q1	0.08-14.76	3.21	3.16	2.43-35.83	13.62	8.30
		168	Q2	0.01-4.22	0.86	0.82	0.46-17.77	5.24	3.67
		168	Q3	0-2.67	0.39	0.43	0.11-11.09	2.72	2.11
		168	Q4	0-1.41	0.14	0.21	0.03-6.12	1.07	1.06
		168	**Overall**	**0.03-5.18**	**1.19**	**1.10**	**0.78-16.34**	**5.83**	**3.74**

We also looked at average percentage marks lost from all unattempted items for individual ability groups taking the papers and Figure A2 shows their distributions. Some of the summary statistics for the subjects are also listed in Table A1. Although the patterns shown in Figure A2 are similar to those seen in Figure A1, the magnitudes are noticeably larger. For the higher tier papers, except for those from combined science, average percentage marks associated with all unattempted items are below 1% of maximum paper marks for the highest ability group (Q4). For combined science, these vary from 0.2% to 2.4%. For the foundation tier papers, average percentage marks lost for the highest ability group (Q4) are close to 1% of max paper marks for biology, chemistry and physics and well above 1% for maths and combined science. For geography, average percentage marks lost from all unattempted items are below 1% of maximum paper marks for the highest ability group (Q4). For all the subjects, average percentage marks lost for group Q3 are close to or above 1%. For Q2, these are well above 1%.

It is likely that time pressure will be one of the many factors that can affect responses to all items in a paper, including those in the middle of the paper. The distributions of average percentage marks lost from all unattempted items shown in Figure A2 further suggest that there is a degree of speededness in most of the papers analysed, particularly in the maths and combined science foundation tier papers.

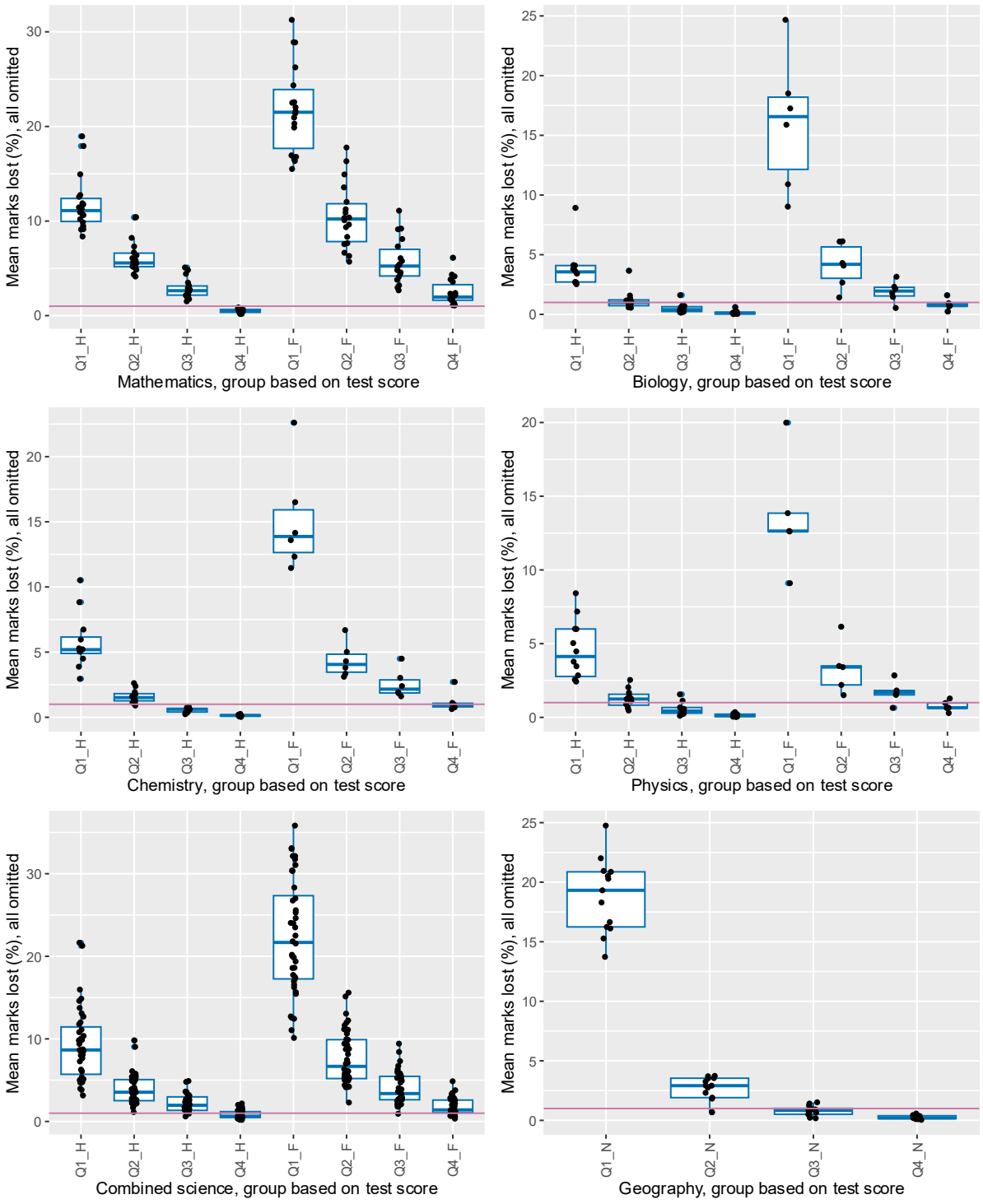

Figure A2 Average percentage marks lost from **all unattempted items** for each of the ability groups taking the papers from the 6 subjects investigated. A horizontal line at y=1.0 is drawn in the graphs

Appendix B: Additional table

Table B1 Values of model parameters which are significantly different from zero at $p<0.05$ for variables included in the multinomial logistic regression analysis for classifications based on the partial credit Rasch model for the 181 papers investigated. The least-speeded group was used as the reference group.

Subject	Spec	Tier	Paper	Group	Abi	G: F	FSM: Y	SEN: Y	Eth: OTH	Lang: OTH	Abi*G (F)	Abi*FSM (Y)	Abi*SEN (Y)	Abi*Eth (OTH)	Abi*Lang (OTH)
Maths	SP1	H	17P1	Less	-2.77	-0.39	0.31	0.24	-0.22		0.34	0.45	0.42		0.29
				Most	-0.65	-0.80	0.31	0.33	-0.40		0.26		0.21	-0.22	
			18P1	Less	-3.33	-0.28	0.42	0.28	-0.35		0.51	0.44	0.33		
				Most	-0.83	-0.64	0.37	0.37	-0.41	-0.20	0.33	0.45	0.27		
			19P1	Less	-3.38		0.58	0.40			0.33	0.47		0.40	
				Most	-0.92	-0.53	0.34	0.46	-0.39		0.15	0.34	0.30		
			17P2	Less	-2.92	-0.10	0.45	0.30			0.43	0.66	0.49	0.20	0.24
				Most	-0.63	-0.44	0.13	0.12	-0.27		0.25	0.44			
			18P2	Less	-1.92	-0.27	0.42				0.17	0.43	0.36	0.26	
				Most	-0.36	-0.50	0.21	0.23	-0.40	-0.26	0.16	0.49	0.32		0.13
			19P2	Less	-3.54	-0.21	0.36	0.43			0.34	0.57	0.48	0.32	
				Most	-1.42	-0.62	0.33	0.24	-0.33	-0.37	0.23	0.48	0.20		
			17P3	Most	-1.05	-0.65		0.47							0.39
			18P3	Less	-3.79	-0.25	0.54	0.50			0.44	0.66	0.62	0.24	0.31
				Most	-1.04	-0.62	0.28	0.23	-0.35		0.21	0.42	0.38		
			19P3	Less	-4.23		0.78	1.17	0.72			0.68	0.88	0.52	
				Most	-1.36		0.39	0.26				0.42	0.31		
		F	17P1	Less	-4.03								0.70		0.90
				Most	-1.07	-0.64	0.28	0.36	-0.15	-0.27		0.26	0.83		0.30
			18P1	Most	-1.36	-0.28	0.19	0.38	0.17				0.55		
			19P1	Less	-4.64										
				Most	-1.50	-0.63	0.29	0.49	-0.29	-0.21	-0.11	0.16	0.63		0.24
			17P2	Less	-1.39	-0.42	0.38	0.52	-0.17			0.39	0.83		0.38
				Most	-0.37	-0.81	0.22	0.26	-0.22	-0.20	-0.07	0.19	0.49		0.16
			18P2	Less	-0.69	-0.20	0.16	0.13			-0.07	0.16	0.43		0.13
				Most	-0.63	-0.53	0.30	0.31		-0.15	-0.12	0.33	0.72	0.12	0.20
			19P2	Most	-1.37	-0.66	0.34	0.36		-0.26	-0.07	0.19	0.51		0.19
			17P3	Less	-1.34	-0.39	0.29	0.38		-0.14		0.27	0.67		0.28

				Most	-0.49	-0.56	0.24	0.17	-0.14	-0.25	0.08	0.13	0.39		0.26
			18P3	Less	0.34	-0.38		-0.09		-0.15	-0.11		0.24		
				Most	0.09	-0.46	0.16			-0.29	-0.10	0.24	0.61	0.13	0.21
			19P3	Less	-1.43	-0.42	0.36	0.48	-0.33		-0.14	0.27	0.75		0.48
				Most	-0.39	-0.69	0.29	0.20	-0.17	-0.24	-0.12	0.19	0.61		0.17
	SP2	H	17P1	Most	-0.79	-0.49				0.69					
			17P2	Less	-2.80										-1.10
				Most	-0.85	-0.36		0.55		0.47			0.57		
			17P3	Less	-2.70									-1.15	
				Most	-1.02	-0.24	0.47			0.30		0.51			
			18P1	Less	-2.92	-0.33									
				Most	-0.68	-0.74					0.21				
			18P2	Less	-3.11	-0.40	0.45		-0.30		0.75				
				Most	-1.08	-0.75	0.36	0.36	-0.29	-0.47	0.34				
			18P3	Less	-2.94	-0.54					0.38				
				Most	-1.01	-1.07	0.26	0.35	-0.40		0.21	0.40			
			19P1	Most	-1.90	-0.95	0.46							0.67	
			19P2	Less	-1.00	-0.12					0.20				
				Most	-0.67	-0.22					0.19			-0.40	0.42
			19P3	Less	-2.41	0.37					0.39	0.49	-0.67		
				Most	-0.87	-0.21	0.29	0.30			0.27				
		F	17P1	Less	0.17	-0.13	0.13		0.16	0.32			0.47		0.21
				Most	-0.36	-0.54	0.39	0.19	0.19	0.30	0.24	0.25	0.27		
			17P2	Less	-0.24	-0.35	0.20						0.36		0.19
				Most	0.19	-0.76	0.19		-0.20		0.22	0.19	0.16		0.25
			17P3	Less	-1.24	-0.21	0.34	0.29				0.22	0.66		0.21
				Most	-0.59	-0.78	0.22	0.24					0.50		
			18P1	Less	-0.84	-0.62	0.25			-0.30	-0.16		0.32		
				Most	-0.63	-0.67	0.23	0.20			-0.16		0.46		
			18P2	Less	0.33	-0.53		-0.13		-0.17			0.35		
				Most		-0.60	0.24	0.13		-0.27	-0.14		0.21		
			18P3	Less	-1.56			0.27	0.29	0.30			0.58		0.24
				Most	-1.04	-0.54	0.27				0.16	0.21	0.57		0.34
			19P1	Less	-2.22		0.32	0.84		0.60		0.31	0.90		0.61
				Most	-0.80	-0.43	0.24		0.16			0.23	0.55		0.40
			19P2	Less	-1.18	-0.28	0.29	0.13			-0.10	0.31	0.66		0.30
				Most	-0.66	-0.71	0.25			-0.41		0.14	0.43		0.21
			19P3	Less	-0.39	-0.43	0.16		-0.18	-0.24	-0.14				

				Most	-0.66	-0.51	0.39	0.34	-0.22	-0.33		0.13	0.42		0.20
Bio	SP1	H	18P1	Most	-1.16	-0.15	0.27	0.23	0.33	0.25	0.16				
			19P1	Most	-1.74	-0.32	0.38	0.54	-0.26			0.26			
			18P2	Most	-2.27	-0.39		0.72		0.31			0.41		0.33
			19P2	Most	-1.91		0.31	0.73	0.27	0.27	0.29		0.32		
		F	18P1	Most	-1.82	-0.64		0.96	0.38				0.63		
			19P1	Most	-1.94	-0.47		0.63	0.59				0.44		0.39
			18P2	Most	-1.59	-0.53							0.41		
			19P2	Less	-1.79	-0.19					0.17		0.37	-0.29	
				Most	-1.72	-0.68	0.48	0.54				0.38	0.97	-0.26	0.29
	SP2	H	18P1	Most	-0.65				0.58						
			18P2	Most	-2.28								0.91		
			19P1	Most	-1.65		0.67					1.25			
			19P2	Most	-2.80		1.15								
		F	19P1	Less	-2.90			0.70					1.09		
				Most		-0.46					0.64				-1.04
			19P2	Less	-2.71			1.06					1.55		
				Most	-1.47	-0.51			0.91					0.81	
	SP3	H	18P1	Most	-1.83			1.70							
			18P2	Less	-2.94										
				Most	-1.52										
			19P1	Most	-3.51		2.34	2.30							
			19P2	Most	-1.27	-0.50	0.90		-0.99						
Chem	SP1	H	18P1	Most	-1.62	-0.48	0.38	0.31							
			19P1	Most	-1.59	-0.45		0.29			0.13				
			18P2	Most	-1.42	-0.43	0.26	0.40			0.11				0.17
				Most	-1.71	-0.28									
		F	18P1	Most	-1.77	-0.61		0.66	-0.36				0.76		0.43
			19P1	Most	-1.15	-0.31							0.25		
			18P2	Most	-1.61	-0.52	0.26	0.35	-0.37			0.30	0.58		
			19P2	Most	-1.81	-1.05	0.49				-0.31				
	SP2	H	18P1	Most	-1.82	-0.64		0.74					0.75		
			18P2	Most	-1.68			0.61	0.57				0.51	0.43	
			19P1	Most	-1.17	-0.55		0.60	0.41					0.45	
			19P2	Most	-2.16	-0.65		0.94							
		F	19P1	Most	-2.36										1.00
			19P2	Most	-2.33								0.93		
	SP3	H	18P1	Most	-1.86										

			18P2	Most	-2.90			1.96					2.53		
			19P1	Most	-4.48										
			19P2	Most	-1.75										
Phy	SP1	H	18P1	Most	-1.15	-0.27		0.45			0.19		0.44		
			19P1	Most	-1.73	-0.20	0.38	0.37					0.49	-0.28	
			18P2	Most	-1.71	-0.40		0.26						0.22	
			19P2	Most	-1.80	-0.83		0.75			-0.20				
		F	18P1	Most	-1.77	-1.06							0.48		0.62
			19P1	Most	-1.25	-0.60	0.44	0.48			-0.33		0.38		0.37
			19P2	Most	-1.85	-0.26		0.53					0.63		0.40
	SP2	H	18P1	Most	-2.35		0.77						0.79		
			18P2	Most	-1.85		1.07					1.06			
			19P1	Most	-3.19		1.89				1.46	2.95			
			19P2	Most	-2.60			1.48							
		F	19P1	Less	-2.68									1.39	
				Most	-1.72	-0.58									
			19P2	Less	-3.39										
				Most	-1.99										
	SP3	H	18P1	Most	-1.78										
			18P2	Most	-2.11			1.69							
			19P1	Most	-1.62										
			19P2	Most	-2.31										
Comb. Sci.	SP1	H	18P1	Less	-5.53								2.86		
				Most	-0.94										
			19P1	Most	-0.84			1.11							
			18P2	Less	-5.33										
				Most	-1.04										
			19P2	Most	-1.81										
			18P3	Less	-1.89										
				Most	-1.23										
			19P3	Less	-3.25					-74.41					-47.43
				Most	-1.39		0.85								
			18P4	Less	-1.10	-0.51			-0.63						
				Most	-0.77										
			19P4	Less	-2.80										
				Most	-0.85	-0.92									
		F	18P1	Less	-5.25					2.39				-1.97	1.76
				Most	-1.86	-1.37		0.55					0.62	-0.93	

			19P1	Less	-2.61	0.59					0.72		0.56		
				Most	-1.47		0.34						0.78		
			18P2	Most	-1.25	-1.02							0.38		
			19P2	Less	-4.55			1.07					1.73		
				Most	-1.77			0.28					1.09		
			18P3	Less	-2.91	-0.57		0.45					1.04		
				Most	-1.63	-0.99	0.38					0.46	0.69	-0.59	
			19P3	Less	-1.14	-0.38	0.35	0.36				0.38	0.96		
				Most		-0.58						0.56	0.51		
			18P4	Less	-3.75	-0.69		0.89					1.27		
				Most	-1.32	-1.29	0.46			-0.94		0.46	0.88		
			19P4	Less	-4.50			0.51			0.75	0.60	1.28		
				Most	-1.80	-0.89			-0.43		0.42		1.41		
			18P1	Most	-1.26	-0.38	0.26	0.22	0.22	0.16					
			19P1	Most	-0.19	-0.38	0.20	0.34			-0.13			-0.21	
			18P2	Most	-1.53	-0.40	0.27	0.37	0.14			0.13	0.31	0.11	
			19P2	Most	-1.25	-0.54	0.26	0.49	0.17	0.26	0.25		0.21		0.18
			18P1	Less	-1.85	-0.37	0.24	0.14				0.22	0.44		
				Most	0.11	-0.48	0.14	0.12	-0.08				0.16	-0.14	
		H	19P1	Most	-1.53	-0.53	0.22	0.37			0.14				
			18P2	Most	-1.14	-0.44	0.24	0.12	-0.09				0.20		
			19P2	Most	-1.35	-0.54	0.28	0.15	-0.16						
			18P1	Most	-1.76	-0.42	0.40	0.16	-0.12			0.21	0.34		
			19P1	Most	-0.97	-0.32	0.22	0.39	0.12	0.18			0.29		
			18P2	Most	-1.70	-0.22	0.26	0.16				0.19			
	SP2		19P2	Most	-1.49	-0.48	0.22	0.20	0.11		0.14				
			18P1	Most	-1.76	-0.52	0.19	0.29	0.17	0.27		0.12	0.34	0.19	
			19P1	Less	-1.20	-0.31			0.17		0.15		0.27	0.15	
				Most	-0.88	-0.52	0.15	0.13	0.10	0.10	0.13	0.09	0.29	0.10	0.09
			18P2	Less	-3.08	-0.58	0.27	0.45		0.49		0.18	0.70		0.54
				Most	-1.35	-0.64	0.24	0.09	-0.20	-0.13		0.24	0.61		0.30
		F	19P2	Less	-3.22	-0.51	0.33	0.72		0.44		0.28	0.72		0.46
				Most	-1.29	-0.64	0.26	0.13	-0.12	-0.09	0.09	0.27	0.53		0.24
			18P1	Less	-2.34	-0.62	0.19	0.17		0.24		0.21	0.53		0.17
				Most	-0.52	-0.80	0.29	0.15		0.06	-0.09	0.19	0.43		
			19P1	Less	-3.41	-0.47	0.39	0.69	-0.17	0.59		0.33	1.01		0.67
				Most	-0.56	-0.49	0.20				-0.07	0.30	0.64		0.27
			18P2	Less	-3.39	-0.55	0.48	0.72	-0.23	0.39		0.41	1.00	-0.14	0.41

				Most	-1.19	-0.56	0.26		-0.18	-0.10		0.32	0.64		0.20
			19P2	Less	-3.88	-0.49	0.39	0.68		0.48		0.35	0.99		0.65
				Most	-1.33	-0.60	0.29	0.05	-0.20	-0.07		0.37	0.81	-0.08	0.32
			18P1	Less	-1.90	-0.20	0.29	0.33				0.31	0.57	-0.13	0.30
				Most	-0.92	-0.23	0.07			-0.18	0.07	0.12	0.45		0.13
			19P1	Less	-3.70	-0.54	0.41	0.70	-0.22	0.57	-0.14	0.43	1.05	-0.32	0.63
				Most	-0.88	-0.57	0.23	0.08				0.40	0.84	-0.10	0.26
			18P2	Less	-3.90	-0.91	0.54	0.53		0.37	-0.34	0.54	0.77		0.34
				Most	-1.52	-0.60	0.23	-0.11	-0.17			0.32	0.59	-0.12	0.21
			19P2	Less	-3.75	-1.14			-0.62		-0.36		0.45	-0.27	0.35
				Most	-1.49	-0.67	0.19			-0.09	-0.07	0.17	0.36		0.16
	SP3	H	18P1	Less	-3.01										
				Most	-1.27	-0.31		0.33	0.42						
			18P2	Most	-1.54	-0.49		0.76							
			18P3	Most	-1.17	-0.44			0.41	0.51					
			18P4	Less	-2.05	-0.70						0.58			
				Most	-0.70	-0.60	0.28	0.47	-0.23		0.22				
			18P5	Most	-1.96	-0.79		0.80					0.62		
			18P6	Less	-3.77			0.83					1.01		
				Most	-1.56	-0.68					0.35				
			19P1	Most	-1.47	-0.73	0.66	0.91							
			19P2	Most	-1.93		0.52	0.73	0.52				0.98		
			19P3	Most	-0.90	-0.49									
			19P4	Most	-1.47				0.45						
			19P5	Less	-3.05	-0.58									
				Most	-0.96	-0.47									
			19P6	Most	-2.27		0.42	0.50	0.38		0.61				
		F	18P1	Less	-1.47	-0.30	0.36					0.28	0.33		
				Most	-1.13	-0.43								0.34	-0.42
			18P2	Less	-1.17	-0.47				0.24		0.36	0.57		
				Most	-1.92	-0.95	0.43					0.47	0.74		
			18P3	Less	-2.64			0.26			0.27	0.34	0.81		0.58
				Most	-1.22	-0.41	0.28			-0.28	0.34		0.61		
			18P4	Less		-0.18							0.22		
				Most	-1.41	-0.72	0.36		-0.38				0.46		0.32
			18P5	Less	-2.27					0.42			0.71		
				Most	-2.03						0.26		0.62		
			18P6	Less	-1.52	-0.25			0.20		-0.23		0.80		0.54

				Most	-0.31	-0.86	0.33	-0.40	-0.22					0.37	
			19P1	Less	-2.01	-0.45		0.26				0.32	0.69		
				Most	-1.00	-0.40	0.35				0.42		0.31	0.46	-0.34
			19P2	Less	-1.88	-0.41						0.28			
				Most	-2.20	-0.58	0.27					0.34	0.78		
			19P3	Less		-0.24									-0.31
				Most	-0.75	-0.50							0.29		
			19P4	Less	-1.79	-0.74							0.39		
				Most	-1.27	-0.85	0.22		-0.22				0.49		
			19P5	Less	-1.41	-0.31							0.45		
				Most	-1.31	-0.74	0.34			-0.27		0.24	0.55		
			19P6	Less	-0.73										
				Most	-1.38	-0.81	0.34		-0.36	-0.33		0.37	0.49		
	SP4	H	18P1	Most	-1.78				0.97					1.10	
			18P2	Most	-1.31										
			18P4	Less	-2.01	-0.66									
				Most	-1.87	-0.84									
			19P1	Less	-2.30							2.05			
				Most		-0.51			0.54			1.78			
			19P2	Most	-1.06					-0.89			3.90		
			19P3	Most	-0.96	-0.53									
			19P4	Most	-4.16										
		F	18P1	Less	-3.24	-1.00						0.91	0.92		
				Most	-1.34	-0.76			0.47						
			18P2	Less	-3.87								1.40		
				Most	-1.46	-0.48							1.04		
			18P3	Less	-4.73			0.94					1.99		
				Most	-0.65	-0.62							0.75		
			18P4	Less	-3.92				1.04				1.18	1.25	
				Most	-1.18	-1.07									
			19P1	Less	-1.47	-0.58							0.81		
				Most		-0.63	0.37								
			19P2	Less	-5.54										
				Most	-1.58	-0.56							0.92	-0.69	0.79
			19P3	Most	-2.89	-1.47					-0.95	0.89	1.27		
			19P4	Less	-4.72					1.80			1.79	-1.82	2.24
				Most	-2.28								1.28		
Geog	SP1		19P3	Less	-3.41	-0.22	0.45	0.84				0.31	0.92		0.16

				Most	-1.78	-0.42	0.44	0.53		0.11	0.06	0.25	0.77		0.19
	SP2		18P1	Most	-1.69								0.61		
			18P2	Less	-3.30										
				Most	-1.52	-0.69	0.67	0.66				0.35	0.63		
			18P3	Less	-4.26		0.98	1.38				0.93	1.45		
				Most	-2.54	-0.71	0.58	0.75				0.81	1.22		0.47
			19P1	Most	-1.92		0.54						0.42		
			19P2	Less	-3.72		1.40	1.72				0.74	1.49		
				Most	-1.63	-0.28	0.74	0.51				0.39	0.88		0.39
			19P3	Less	-3.45			1.45					1.07		
				Most	-2.12	-0.28	0.55	0.74			0.35		1.17		
	SP3		18P1	Most	-2.25	-0.79	0.49	0.70				0.32	0.80		
			18P2	Most	-2.58	-0.85		0.41					0.59		0.44
			18P3	Most	-2.97	-0.40		0.90					0.90		
			19P1	Most	-2.37	-0.78	0.43	0.58	0.28			0.31	0.72		0.47
			19P2	Most	-2.28	-0.74		0.70					0.79		
			19P3	Most	-2.77	-0.51		0.83					0.97		0.39

Published by:

ofqual

2nd floor, 1 Friargate
Station Square
Coventry
CV1 2GN

0300 303 3344

www.gov.uk/ofqual

November 2025

Ofqual/25/7267/3