Research Matters / 39
A Cambridge University Press & Assessment publication

ISSN: 1755-6031

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

Accessibility of GCSE science questions that ask students to create and augment visuals: Evidence from question omit rates

Santi Lestari

To cite this article: Lestari, S. (2025). Accessibility of GCSE science questions that ask students to create and augment visuals: Evidence from question omit rates. *Research Matters: A Cambridge University Press & Assessment publication, 39,* 39–65. https://doi.org/10.17863/CAM.116169

To link this article: https://www.cambridgeassessment.org.uk/Images/research-matters-39-accessibility-of-gcse-science-questions-that-ask-students-to-create-and-augment-visuals.pdf

Abstract:

The ability to draw visual representations such as diagrams and graphs is considered fundamental to science learning. Science exams therefore often include questions which require students to draw a visual representation, or to augment a partially provided one. The design features of such questions (e.g., layout of diagrams, amount of answer space) could, however, influence students' ability to respond to the questions and present potential accessibility issues, which in turn could influence the validity of score inferences. This article reports on a small-scale study examining the accessibility of GCSE science questions involving the creation and augmentation of visuals (e.g., adding an element to a partially provided diagram) by analysing the patterns of question omit rates. Omit rates for questions involving creating or augmenting visuals were compared to those for questions without, and these comparisons were conducted across tiers, subjects, question position, maximum marks and facility values, as well as by gender and attainment group.

Accessibility of GCSE science questions that ask students to create and augment visuals: Evidence from question omit rates

Santi Lestari (Research Division)

Introduction

Visual representations including graphs, diagrams, images and illustrations are prevalent in science texts and play a key role in science communication (Trumbo, 1999). They are often used to support verbal descriptions or explanation of complex scientific concepts and processes (Wang & Wei, 2024). Scientific visual literacy has therefore received considerable attention in science education and has been a feature in science education reform in several jurisdictions (LaDue et al., 2015; Wang & Wei, 2024). Scientific visual literacy encompasses not only the ability to interpret scientific visual representations but also to create them. There are ample arguments for, and evidence of, how visual representation construction is core to science learning (e.g., Ainsworth et al., 2011; Prain & Tytler, 2012; Tytler et al., 2018; Tytler et al., 2020). Therefore, including questions which require students to create visual representations in exams has been strongly advocated (Unsworth & Herrington, 2023; Wang & Wei, 2024). Chang et al. (2020) also note that drawing is a powerful method to assess students' understanding of scientific concepts and emphasise the advantages of requiring students to draw, rather than to write, for assessing certain concepts in science. As such, the Department for Education GCSE Science subject content document includes not only interpreting data presented in visual forms, but also communicating scientific observations and concepts through the creation of visual representations, as skills to be developed and assessed under "working scientifically" (Department for Education, 2015).

Exam question features could affect students' ability to engage with an exam question, i.e., to understand the question and subsequently to respond to it to demonstrate their knowledge, skills and understanding (Crisp & Macinska, 2020). For exam questions which require students to create a visual representation or augment a partially provided one, question features that could potentially influence students' performance include the layout of the visual representation and the amount of answer space. For example, if the answer space is too restricted due to certain layout formatting of the question, students might struggle to fit their answer within the space and therefore their ability to demonstrate their understanding could be compromised. In short, design features of an exam or exam question could present accessibility issues which in turn may weaken the validity of test score inferences (Beddow, 2012).

Given the important influence of accessibility on validity, awarding organisations are required to ensure accessibility of their exams (Ofqual, 2022). OCR has a set of accessibility principles for GCSE Science to provide guidance in test construction and, thus, to ensure that all students can demonstrate their knowledge, skills and understanding (OCR, 2018a, 2018b). Some of these principles specifically relate to the use of visuals (i.e., the inclusion, placement and layout of visuals), and one principle in particular concerns questions which require students to do something with a visual (i.e., the visual will be centred with sufficient space around it to allow students to fit in their response). Such principles can help ensure that questions are as accessible as possible for candidates, but there is also a role for ongoing evaluation of the accessibility of exam questions.

There are multiple ways to investigate the accessibility of exams and exam questions. One method is by collecting expert judgements. For example, Beddow et al. (2013) asked test development experts to review exam questions using the *Accessibility Rating Matrix*. One of the elements assessed in the matrix is the use of visuals (e.g., the complexity of visuals and the placement of visuals). Another method involves gathering students' perspectives. For example, Crisp and Macinska (2020) interviewed students to gather their perspectives on the accessibility of GCSE Science questions. Other methods could involve conducting linguistic analyses of exam questions, as exemplified by Beauchamp and Constantinou (2020).

In this article, we argue that analyses of question omit rates could provide information about question accessibility. The omit rate for a question refers to the proportion of students who did not provide a response. Given that general qualifications in England use positive marking and, thus, there is no penalty for providing an incorrect answer, it is in the candidate's best interest to try to answer all questions (Sarac & Loken, 2023).

While research mostly focuses on the quality of student responses in an exam (i.e., correct, partially correct and incorrect responses), omit rates could also provide additional information about the exam (Papanastasiou, 2020) and could be useful to investigate various aspects of exams such as speededness (e.g., Walland, 2024) and differential test functioning (e.g., Ben-Shakhar & Sinai, 1991). Omit rates, however, have not been commonly used to investigate exam accessibility even though they could be an indicator of accessibility barriers. If certain questions or question types have systematically high omit rates, this could indicate potential access barriers. It could be argued that the nature and level of the demands of questions also contributes to variability in question omit rates. Referring to the CRAS scale of demands[1] (Pollitt et al., 2007), questions that require students to create a visual or augment a partially provided one can be considered to have a distinct and potentially higher level of strategy demand. In particular, response strategy demand, whereby students are required to organise how to

1 CRAS stands for Complexity, Resources, Abstractness and Strategy.

communicate their response through a visual representation, could be affected. Performance data (i.e., correct, partially correct and incorrect responses) may mask these potential accessibility barriers.

While the current research explores the use of omit rates as a possible indicator of accessibility, it is important to be aware that various factors could contribute to questions being omitted. Previous research has shown that omit rates can be influenced by characteristics of the student (e.g., ability level, gender, cultural background), characteristics of the exam and exam question (e.g., exam content, question format, question difficulty, question position) and interactions of the two. Examining the pattern of question omit rates in a low-stakes multiple-choice reading comprehension assessment, Clemens et al. (2015) found that students from lower performing subgroups had higher omit rates than those from higher performing subgroups, especially on the questions towards the end of the test.

In a larger-scale study involving high-stakes GCSE exams in biology, chemistry, physics, science and mathematics with mixed question formats, Walland (2024) found that omit rates for questions towards the end of exam papers were much higher for students from the lowest achieving subgroup than for those from the other subgroups. The foundation tier papers also had higher omit rates for questions towards the end of the papers than the higher tier papers. While skipping difficult questions could be an indicator of students' use of test-taking strategies, higher omit rates for questions towards the end of the test and especially for lower attaining students are more indicative of this particular group of students not being able to finish the test. This could be because lower attaining students might tend to take more time to attempt questions more generally, including those presented earlier in the test, as they find them harder than their higher attaining peers would. This would result in lower attaining students being more likely to leave questions towards the end of the test unanswered. In addition, it could also be the case that lower attaining students do not have sufficient knowledge, skills and understanding to be able to make an attempt at these later questions, given the tendency for a rough progression of question difficulty through a paper. Therefore, the patterns of question omit rates identified in Clemens et al. (2015) and Walland (2024) seem more likely to be due to the interaction between students' ability level and the difficulty of the subject content in the questions than due to accessibility issues.

Students' gender has also been found to interact with omit rates. Male students were typically found to omit fewer questions than female students in multiple-choice tests (e.g., Ben-Shakhar & Sinai, 1991). However, these patterns of omit rates across genders could vary across question formats and subjects. Matters and Burnett (1999), researching the high-stakes Queensland Core Skills Test, found that for multiple-choice questions omit rates in general were very small and the difference across gender categories was negligible. For constructed response questions, however, omit rates were higher and male students omitted more questions than female students. In von Schrader and Ansley's (2006) analysis of the high-stakes Iowa Tests of Basic Skills and Iowa Tests of Educational Development, female students tended to omit more questions in the mathematics exam while male students tended to omit more questions in the reading and vocabulary exams.

The research reported in the current article examined the accessibility of GCSE Science questions involving the creation and augmentation of visuals (e.g., adding an element to a partially provided diagram) by analysing the patterns of question omit rates.

The main question guiding the research was:

- Is there empirical evidence of atypically high omit rates for GCSE Science items that require diagram creation or augmentation, which could indicate a potential accessibility issue?

To address the research question, omit rates for questions involving creating or augmenting visuals were compared to those for questions without, and these comparisons were conducted across tiers (i.e., foundation and higher tiers), subjects (i.e., biology, physics, chemistry and combined science), question position within a paper, maximum marks and facility values, as well as by gender and attainment group. Comparisons across different question attributes and candidate characteristics were made to help differentiate factors other than question accessibility that may have contributed to omit rates.

Method

Data

In this research, we used the item-level data (marks or omission information for each candidate on each item) from eight OCR GCSE Science specifications[2] from the June 2023 exam series. Each specification had different numbers of papers, and in total there were 44 papers.[3]

Approach to item categorisation

While visual representations are often used in GCSE Science exams, in this study we specifically focused on questions which require students to create a visual representation or augment a partially provided one (e.g., drawing a line of best fit on a graph or completing a diagram). For the sake of brevity, such questions are referred to as "items with diagram(s)" in the remainder of this article.

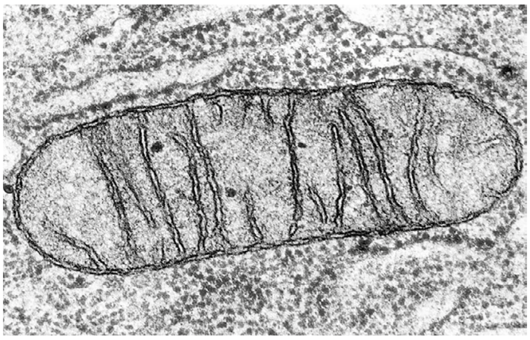
To illustrate, Item b(i) in Figure 1 is an example of an item with a diagram because it requires students to create a scientific drawing. Items 19a(i) and (ii) in Figure 2 are also both examples of items with diagrams because they require students to augment a partially provided graph. Conversely, although Item b in Figure 3 is based on a diagram, it is not considered an item with a diagram in this

2 The eight specifications were: Science A, Combined (9–1) – Gateway Science Suite J250; Biology A (9–1) – Gateway Science Suite J247; Chemistry A (9–1) – Gateway Science Suite J248; Physics A (9–1) – Gateway Science Suite J249; Science B, Combined (9–1) – Twenty First Century Science Suite J260; Biology B (9–1) – Twenty First Century Science Suite J257; Chemistry B (9–1) – Twenty First Century Science Suite J258; Physics B (9–1) – Twenty First Century Science Suite J259.

3 The data used in this research was collected as part of the usual marking and processing of candidates' examination scripts. Data has been stored and used in line with Cambridge University Press & Assessment's Data Privacy notice (https://www.cambridge.org/legal/candidate-privacy-notice).

study because it requires students to explain a process rather than creating or augmenting a diagram.

(b) The image is of a mitochondrion.



(i) Draw the mitochondrion in the box. Your drawing should be a scientific drawing.

[2]

Figure 1: A sample item requiring students to create a visual representation (categorised as "item with diagram(s)")[4]

4 Source: https://www.ocr.org.uk/Images/704945-question-paper-paper-1.pdf

19 A student investigates the effect of pH on an enzyme called catalase.
Catalase breaks down hydrogen peroxide into water and oxygen.

The student collects the oxygen produced by the reaction.
The table shows their results.

pH	Volume of oxygen collected (cm^3)
2	1
4	12
6	24
8	26
10	8

(a) **(i)** Plot a graph of the results. **[2]**

(ii) Draw a line of best fit. **[1]**

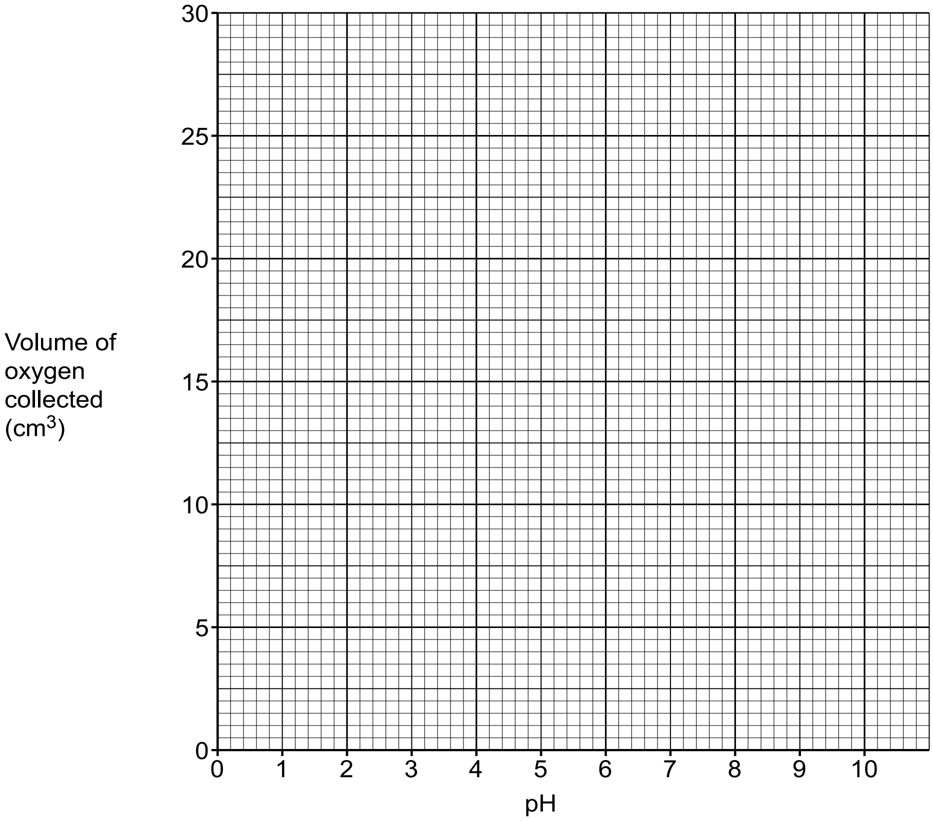
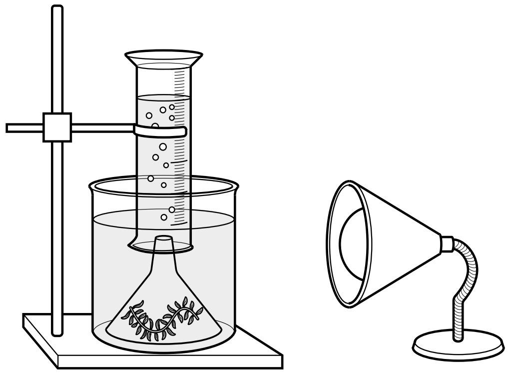


Figure 2: Sample items requiring students to augment a partially provided visual representation (both items categorised as "item with diagram(s)")[5]

5 Source: https://www.ocr.org.uk/Images/704945-question-paper-paper-1.pdf

(b) The student sets up a second experiment using the equipment in **Fig. 3.2**.

Fig. 3.2



Explain how this second experiment will improve the quality of the data collected to measure the rate of photosynthesis.

....

....

....

.... **[2]**

Figure 3: A sample item requiring students to explain a scientific process illustrated in a diagram (categorised as "item without diagrams")[6]

Procedure

The item-level data from the exam papers were processed in three data preparation steps:

1. **Coding of items.**
   As described and exemplified in the previous section, items were binary coded as **"item with diagram(s)"** or **"item without diagrams"**. It should be noted that each paper had only small numbers of items with diagrams (typically three to four), and of 44 papers, only six had more than five items with diagrams.

2. **Removal of multiple-choice item data.**
   Initial exploration of the data suggested that multiple-choice items tended to have zero or close to zero omit rates, which is unsurprising considering the possibility of guessing and the question position (first section of the paper). These very low omit rates would skew the omit rate distribution for a whole paper. Therefore, the multiple-choice item data were excluded from the analysis.[7]

6 Source: https://www.ocr.org.uk/Images/705023-question-paper-combined-science.pdf
7 Only J247, J248, J249 and J250 specifications had multiple-choice items. There were 10 multiple-choice items in each paper in J250, and 15 multiple-choice items in each paper in J247, J248, and J249.

3. **Calculations based on item-level data.**
   Firstly, the **omit rate** needed to be calculated for each item in each paper. Omit rate refers to the proportion of students who did not attempt an item. The value ranges from 0 to 1, with 0 indicating that all students attempted the item and 1 indicating that no student attempted the item.

   To further examine the patterns of omit rates for different gender and attainment groups, disaggregated omit rates also had to be calculated. Item omit rates for each gender group were calculated.[8] Students were also classified into attainment quartiles based on the total marks they achieved in each paper. Then, the item omit rate for each attainment quartile was calculated.

   As the papers had different numbers of items and maximum marks, **item position** within a paper needed to be standardised. Item position was therefore defined as the proportion of how far through the paper an item was in terms of the paper maximum mark. The value ranged from 0 to 1 and was classified into quintiles (i.e., 0.00–0.20, 0.21–0.40, 0.41–0.60, 0.61–0.80 and 0.81–1.00).

   Item maximum mark in the papers included in this study ranged from 1 to 6. However, there were very few items with a maximum mark of 4 to 6. Therefore, maximum marks of 4 to 6 were grouped together into "4 or above".

   The **facility value** for each item in each paper also needed to be calculated. Facility value refers to the mean mark on the item as a proportion of maximum mark and is a useful measure of item difficulty on exams where all the items are compulsory. The value ranges from 0 to 1, with 0 indicating that no marks were scored by any students and 1 indicating that all students achieved maximum marks on the item. To facilitate analysis, item facility value was classified into quintiles (i.e., 0.00–0.20, 0.21–0.40, 0.41–0.60, 0.61–0.80 and 0.81–1.00).

   Finally, all of the variables calculated based on the item-level data for all the 44 GCSE Science papers were combined into a single dataset for further analyses.

The dataset was analysed using descriptive statistics to address the research question. More specifically, the omit rates for items with diagrams were compared with those for items without diagrams. These comparisons were conducted across:

a. tiers (foundation and higher tiers)
b. subjects (i.e., biology, physics, chemistry and combined science)
c. item positions within a paper
d. item maximum marks
e. item facility values (a measure of item difficulty level)

8 The analysis across gender groups did not include data for candidates with no gender information due to the extremely small size of this group.

We also examined the patterns across students from different gender and attainment groups. Boxplots were used to visualise the results. Both the descriptive statistics analyses and boxplot generation were conducted in RStudio (Posit team, 2023).

Results

The results are presented for each of the five aspects (i.e., omit rates by tier, subject, item position, item maximum mark and item facility value) in turn.

Boxplots are used to visualise the distribution of item omit rates and accompanied by tables showing their associated descriptive statistics. In a boxplot, the horizontal line dividing the box into two represents the median value. The line on the lower edge of the box represents the lower quartile, and the line on the higher edge represents the upper quartile. The lines extending from the box, known as the whiskers, represent the variability in the dataset beyond the lower and upper quartiles. The individual dots represent the outliers.

It is important to remember that the proportion of items with diagrams in each paper was generally very small (8 per cent on average), so the results should be interpreted cautiously. It is recommended to consult the descriptive statistics tables that contain the number of items for each category.

Omit rate by tier

As shown in Figure 4 and Table 1, omit rates were generally higher in the foundation tier papers than in the higher tier papers, for both items with and without diagrams. In fact, omit rates for items in the higher tier papers were all very low on average, making it difficult to examine differences in omit rates across tiers as well as across item types within the higher tier papers. Within the foundation tier, although the median omit rate for items with diagrams appeared slightly lower than that for items without diagrams, this difference was still too small to be meaningful.

Analysis of the disaggregated data by attainment group (quartiles) based on candidate overall performance in each paper showed that for both the foundation and higher tiers, omit rates were higher in the lower attainment groups and decreased in the higher attainment groups (see Figure 5 and Table 2). Omit rates were considerably higher for the lowest attainment group (Q1) in the foundation tier than the other quartiles. This was true for both items with and without diagrams, with no meaningful differences observed. Omit rates in the foundation tier papers for candidates in Q2, Q3 and Q4 were broadly comparable to omit rates in the higher tier papers.

Further analysis of the disaggregated data by gender showed that male candidates in the foundation tier papers generally had a higher propensity to omit items than their female peers did, and this was true for both items with and without diagrams, although the difference might be negligible (see Figure 6 and Table 3).

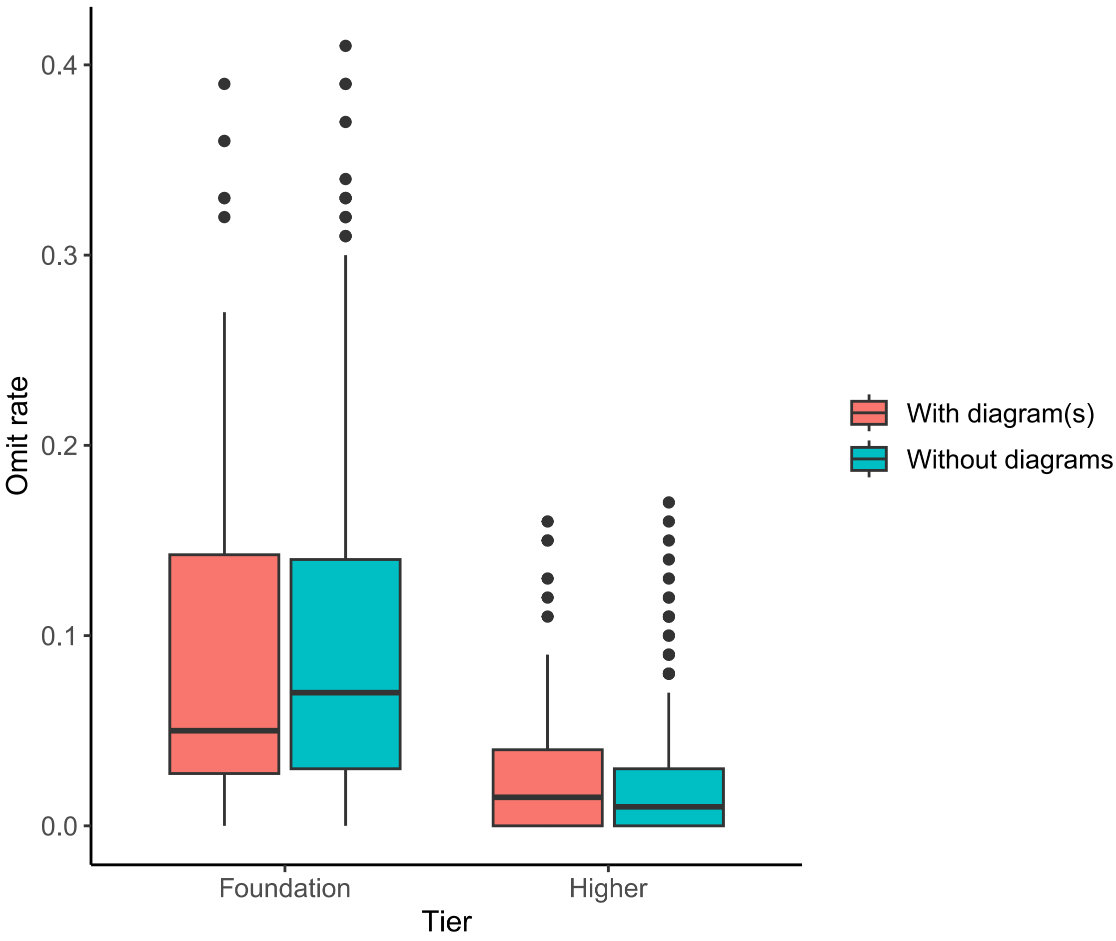


Figure 4: Omit rate by tier

Table 1: Omit rate descriptive statistics, by tier

Tier	Item type	Number of items	Min	Max	Median	Mean	SD
Foundation	With diagram(s)	84	0	0.39	0.05	0.10	0.10
	Without diagrams	793	0	0.41	0.07	0.09	0.08
Higher	With diagram(s)	76	0	0.16	0.01	0.03	0.04
	Without diagrams	718	0	0.17	0.01	0.02	0.02

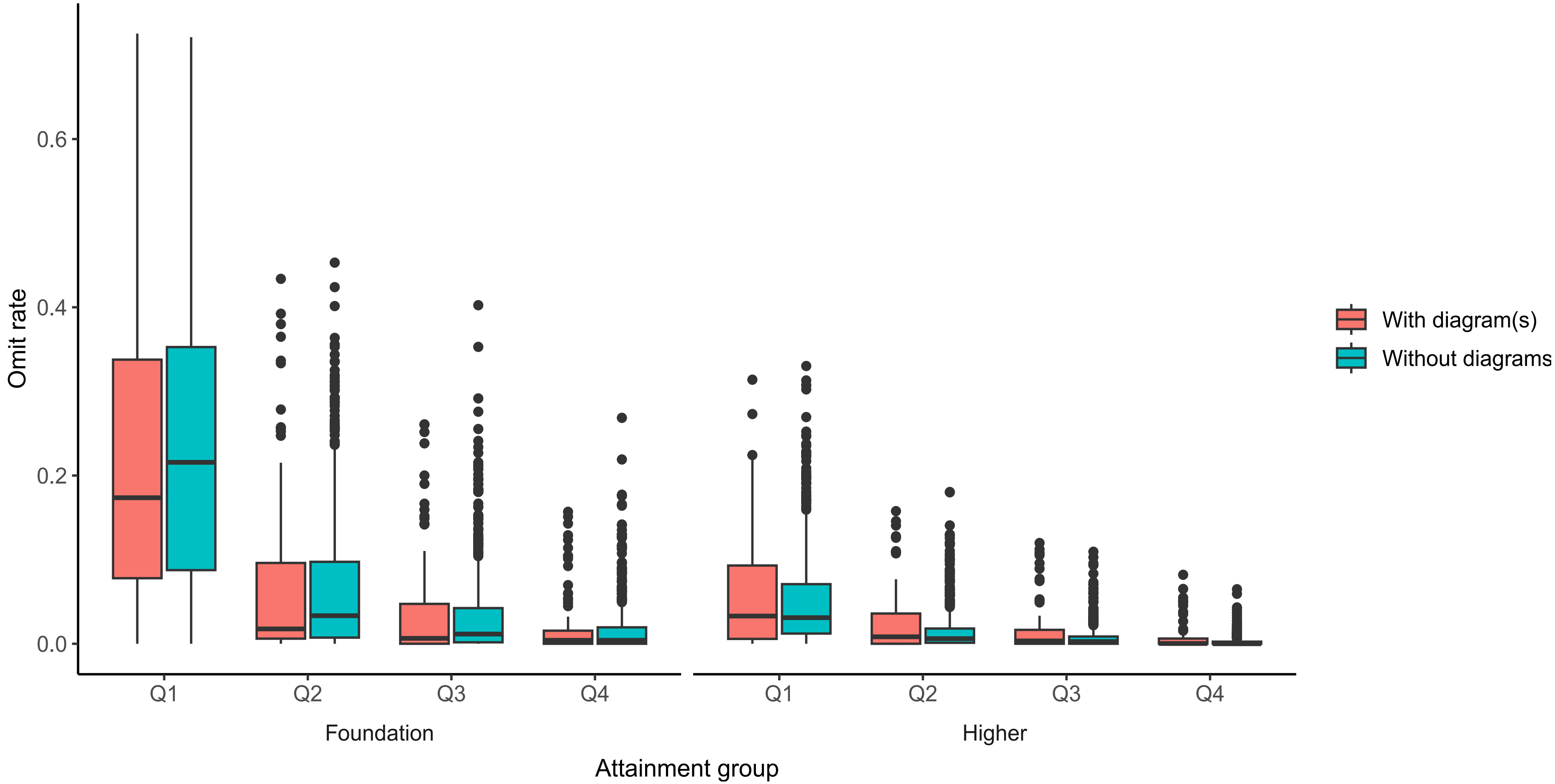


Figure 5: Omit rate by tier and attainment group (Q1 being the lowest attaining quartile and Q4 the highest attaining quartile)

Table 2: Omit rate descriptive statistics, by tier and attainment group (Q1 being the lowest attaining quartile and Q4 the highest attaining quartile)

Tier	Attainment group	Item type	Number of items	Min	Max	Median	Mean	SD
Foundation	Q1	With diagram(s)	84	0	0.73	0.17	0.24	0.20
		Without diagrams	793	0	0.72	0.22	0.24	0.17
	Q2	With diagram(s)	84	0	0.43	0.02	0.08	0.11
		Without diagrams	793	0	0.45	0.03	0.07	0.08
	Q3	With diagram(s)	84	0	0.26	0.01	0.04	0.07
		Without diagrams	793	0	0.40	0.01	0.03	0.05
	Q4	With diagram(s)	84	0	0.16	0	0.02	0.04
		Without diagrams	793	0	0.27	0	0.02	0.03
Higher	Q1	With diagram(s)	76	0	0.31	0.03	0.06	0.07
		Without diagrams	718	0	0.33	0.03	0.05	0.06
	Q2	With diagram(s)	76	0	0.16	0.01	0.03	0.04
		Without diagrams	718	0	0.18	0.01	0.02	0.02
	Q3	With diagram(s)	76	0	0.12	0	0.02	0.03
		Without diagrams	718	0	0.11	0	0.01	0.01
	Q4	With diagram(s)	76	0	0.08	0	0.01	0.02
		Without diagrams	718	0	0.06	0	0	0.01

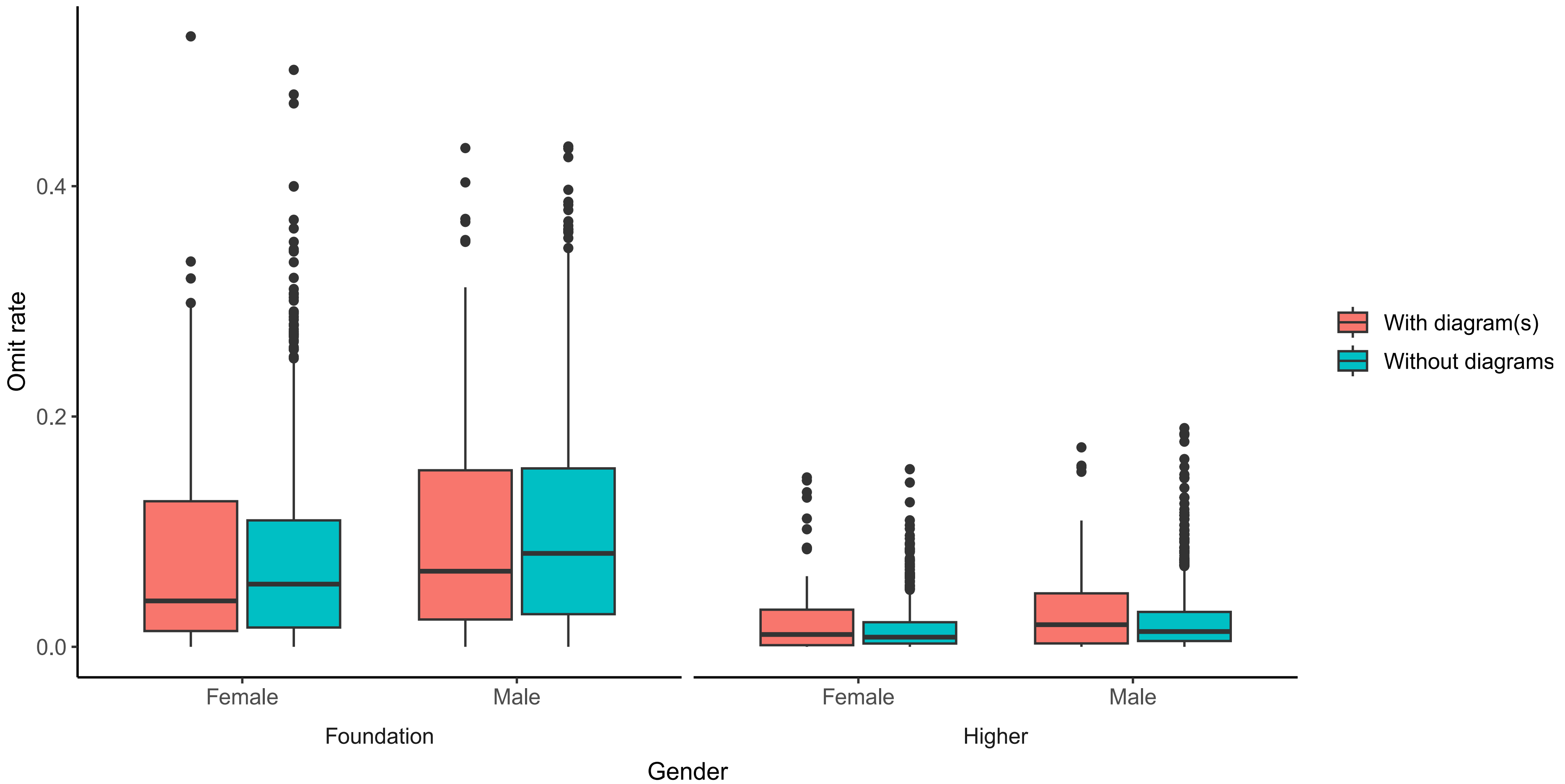


Figure 6: Omit rate by tier and gender

Table 3: Omit rate descriptive statistics, by tier and gender

Tier	Gender	Item type	Number of items	Min	Max	Median	Mean	SD
Foundation	Female	With diagram(s)	84	0	0.53	0.04	0.09	0.10
		Without diagrams	793	0	0.50	0.05	0.08	0.08
	Male	With diagram(s)	84	0	0.43	0.07	0.11	0.11
		Without diagrams	793	0	0.43	0.08	0.10	0.09
Higher	Female	With diagram(s)	76	0	0.15	0.01	0.03	0.04
		Without diagrams	718	0	0.15	0.01	0.02	0.02
	Male	With diagram(s)	76	0	0.17	0.02	0.03	0.04
		Without diagrams	718	0	0.19	0.01	0.02	0.03

Omit rate by subject

Figure 7 and Table 4 show the omit rates by tier and subject. Omit rates varied across papers assessing different subjects. Note that there was only one combined science paper in each tier, hence the small numbers of items in these papers. Chemistry papers overall had higher omit rates than other papers, particularly in the foundation tier. Furthermore, chemistry items with diagrams in the foundation tier appeared to have higher omit rates than those without diagrams. Conversely, biology items with diagrams in the foundation tier tended to have lower omit rates than those without diagrams. There were no substantial differences in omit rates for items with and without diagrams in physics papers. While Figure 7 shows differences in the median omit rates between items with diagrams and items without diagrams in the combined science papers, these differences were not meaningful due to the low numbers of items with diagrams (i.e., only five items in the foundation tier and three items in the higher tier).

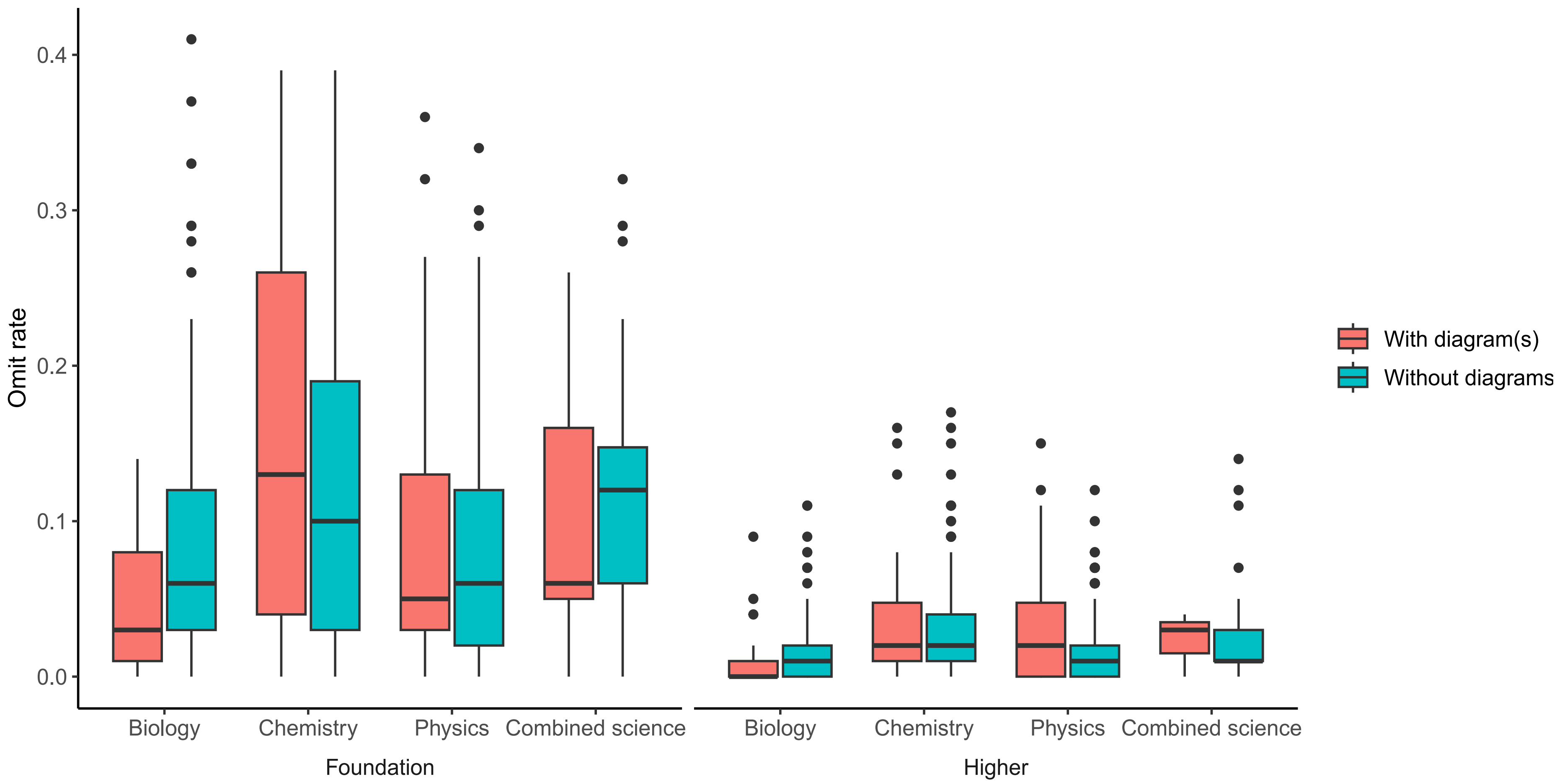


Figure 7: Omit rate by subject

Table 4: Omit rate descriptive statistics, by subject

Tier	Subject	Item type	Number of items	Min	Max	Median	Mean	SD
Foundation	Biology	With diagram(s)	21	0	0.14	0.03	0.04	0.04
		Without diagrams	243	0	0.41	0.06	0.08	0.07
	Chemistry	With diagram(s)	21	0	0.39	0.13	0.16	0.13
		Without diagrams	185	0	0.39	0.10	0.12	0.10
	Physics	With diagram(s)	37	0	0.36	0.05	0.09	0.09
		Without diagrams	333	0	0.34	0.06	0.08	0.07
	Combined science	With diagram(s)	5	0	0.26	0.06	0.11	0.10
		Without diagrams	32	0	0.32	0.12	0.12	0.08
Higher	Biology	With diagram(s)	21	0	0.09	0	0.01	0.02
		Without diagrams	234	0	0.11	0.01	0.01	0.02
	Chemistry	With diagram(s)	22	0	0.16	0.02	0.04	0.05
		Without diagrams	172	0	0.17	0.02	0.03	0.03
	Physics	With diagram(s)	30	0	0.15	0.02	0.03	0.04
		Without diagrams	285	0	0.12	0.01	0.02	0.02
	Combined science	With diagram(s)	3	0	0.04	0.03	0.02	0.02
		Without diagrams	27	0	0.14	0.01	0.03	0.04

Omit rate by item position

Figure 8 and Table 5 show omit rates by tier and item position within a paper. There appeared to be an overall tendency that omit rates increased as the paper progressed. In other words, items towards the end of the paper tended to have slightly higher omit rates than those preceding them. For foundation tier papers, in particular, there was more variability in omit rates for items towards the end of the paper. From the middle to the end of the foundation tier paper, items with diagrams appeared to have lower omit rates than items without diagrams. However, it is hard to be sure that this trend was not random given the low numbers of items with diagrams.

Further analysis of omit rates by item position across different attainment groups (not reported in full here for reasons of brevity) indicated that the rate of omission for items towards the end of the paper was substantially higher for candidates in the lower attaining groups and especially in the foundation tier.

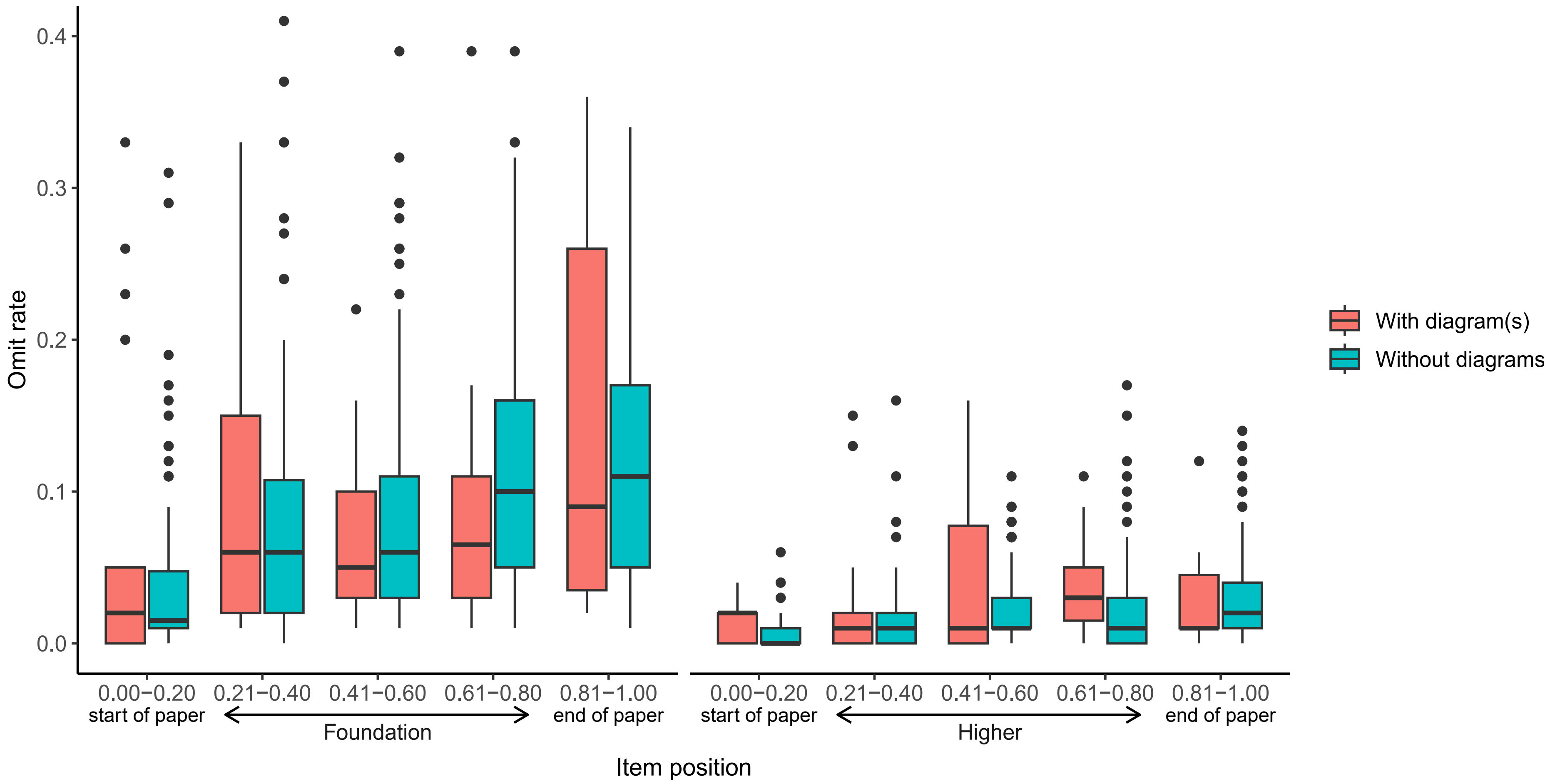


Figure 8: Omit rate by item position

Table 5: Omit rate descriptives, by tier and item position

Tier	Item position	Item type	Number of items	Min	Max	Median	Mean	SD
Foundation	0.00-0.20	With diagram(s)	17	0	0.33	0.02	0.07	0.11
		Without diagrams	102	0	0.31	0.01	0.04	0.06
	0.21-0.40	With diagram(s)	17	0.01	0.33	0.06	0.09	0.09
		Without diagrams	174	0	0.41	0.06	0.08	0.07
	0.41-0.60	With diagram(s)	19	0.01	0.22	0.05	0.07	0.06
		Without diagrams	171	0.01	0.39	0.06	0.08	0.07
	0.61-0.80	With diagram(s)	8	0.01	0.39	0.06	0.11	0.12
		Without diagrams	172	0.01	0.39	0.10	0.11	0.08
	0.81-1.00	With diagram(s)	23	0.02	0.36	0.09	0.14	0.12
		Without diagrams	174	0.01	0.34	0.11	0.12	0.07
Higher	0.00-0.20	With diagram(s)	9	0	0.04	0.02	0.02	0.02
		Without diagrams	104	0	0.06	0	0.01	0.01
	0.21-0.40	With diagram(s)	23	0	0.15	0.01	0.02	0.04
		Without diagrams	168	0	0.16	0.01	0.01	0.02
	0.41-0.60	With diagram(s)	14	0	0.16	0.01	0.04	0.06
		Without diagrams	154	0	0.11	0.01	0.02	0.02
	0.61-0.80	With diagram(s)	15	0	0.11	0.03	0.04	0.03
		Without diagrams	136	0	0.17	0.01	0.02	0.03
	0.81-1.00	With diagram(s)	15	0	0.12	0.01	0.03	0.03
		Without diagrams	156	0	0.14	0.02	0.03	0.03

Omit rate by item maximum mark

Figure 9 and Table 6 show omit rates by tier and item maximum mark. In the foundation tier, it is evident that items with a maximum mark of 1 that had diagrams tended to have higher omit rates than those items with the same maximum mark but without diagrams. Conversely, of items with a maximum mark of 3 and 4 or above, those that had diagrams tended to have lower omit rates than those that did not have diagrams.

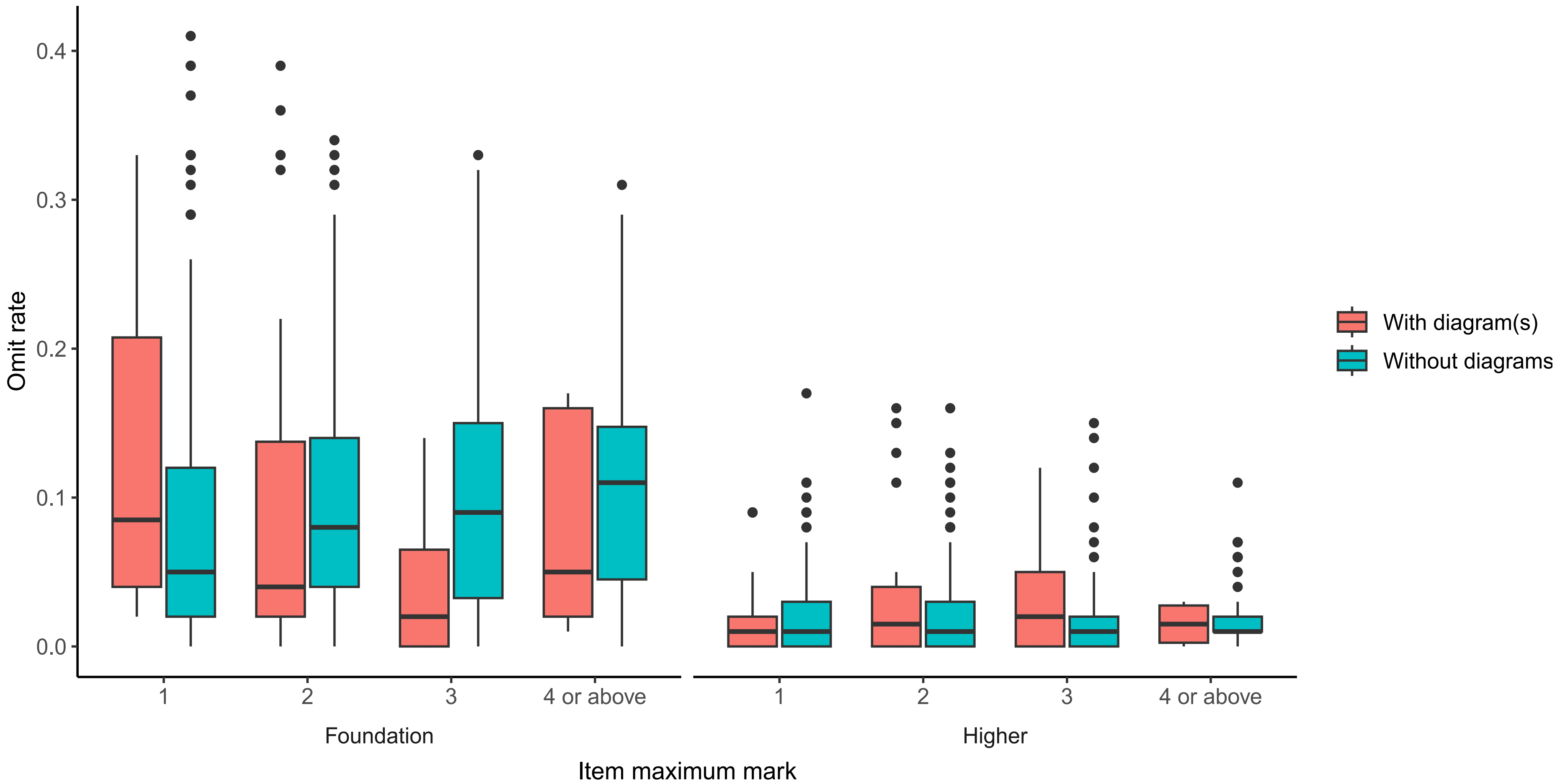


Figure 9: Omit rate by item maximum mark

Table 6: Omit rate descriptive statistics, by tier and maximum mark

Tier	Item max mark	Item type	Number of items	Min	Max	Median	Mean	SD
Foundation	1	With diagram(s)	28	0.02	0.33	0.08	0.12	0.10
		Without diagrams	347	0	0.41	0.05	0.08	0.08
	2	With diagram(s)	40	0	0.39	0.04	0.10	0.12
		Without diagrams	290	0	0.34	0.08	0.10	0.08
	3	With diagram(s)	11	0	0.14	0.02	0.04	0.05
		Without diagrams	110	0	0.33	0.09	0.10	0.08
	4 or above	With diagram(s)	5	0.01	0.17	0.05	0.08	0.08
		Without diagrams	46	0	0.31	0.11	0.11	0.08
Higher	1	With diagram(s)	21	0	0.09	0.01	0.02	0.03
		Without diagrams	251	0	0.17	0.01	0.02	0.02
	2	With diagram(s)	36	0	0.16	0.01	0.03	0.05
		Without diagrams	270	0	0.16	0.01	0.02	0.03
	3	With diagram(s)	13	0	0.12	0.02	0.03	0.04
		Without diagrams	134	0	0.15	0.01	0.02	0.02
	4 or above	With diagram(s)	6	0	0.03	0.01	0.01	0.01
		Without diagrams	63	0	0.11	0.01	0.02	0.02

Omit rate by item facility value

Figure 10 and Table 7 show the comparisons of omit rates between items with diagrams and items without diagrams at similar difficulty levels. Overall omit rates were higher for more difficult items (lower facility values) than for easier items. This is expected because the calculation of facility value also takes into account omissions. The omit rate of an item effectively limits the maximum possible facility value of an item since candidates who did not answer will have received 0 marks on the item (e.g., an item with an omit rate of 0.20, cannot have a facility value of more than 0.80).

It also appeared that items with diagrams tended to have higher omit rates than those at similar difficulty levels but without diagrams. This observation was particularly prominent for items with very low facility values (0.00–0.20), i.e., for very difficult items. While this could be taken to suggest that items with diagrams might have introduced access barriers, it is important to be cautious at interpreting this finding given that there were far fewer items with diagrams than without diagrams at this difficulty level — nine versus 176 in the foundation tier. A closer look showed that all of the nine items with diagrams were located towards the end of the papers. Given that items towards the end of the papers tended to have higher omit rates, it could be that the considerable difference in omit rates between items with diagrams and without was not solely or primarily due to the difficulty level and accessibility, but due to students not being able to reach these items.

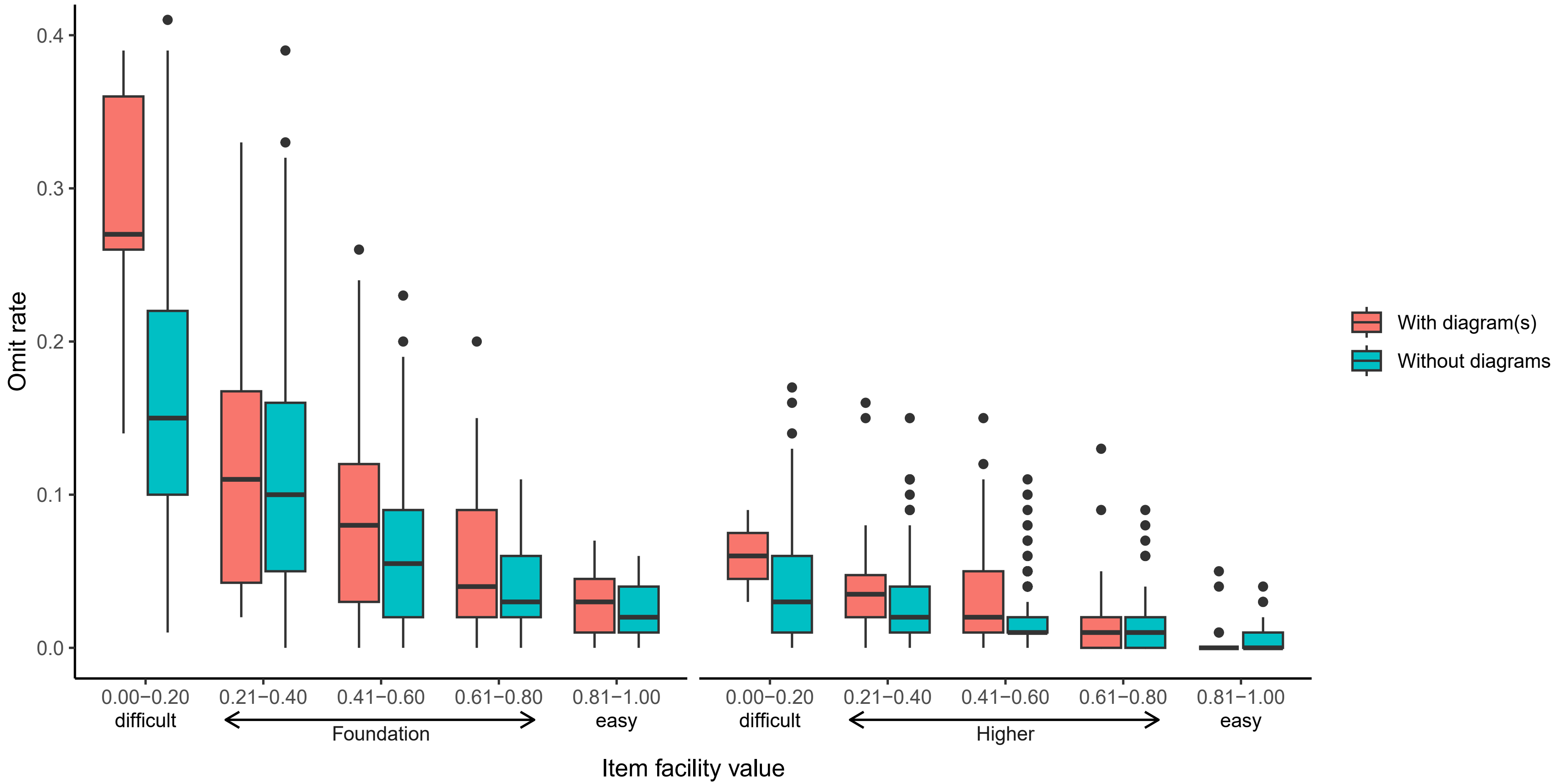


Figure 10: Omit rate by item facility value

Table 7: Omit rate descriptive statistics, by tier and facility value

Tier	Facility value	Item type	Number of items	Min	Max	Median	Mean	SD
Foundation	0.00-0.20	With diagram(s)	9	0.14	0.39	0.27	0.28	0.09
		Without diagrams	176	0.01	0.41	0.15	0.16	0.09
	0.21-0.40	With diagram(s)	14	0.02	0.33	0.11	0.14	0.12
		Without diagrams	217	0	0.39	0.10	0.11	0.07
	0.41-0.60	With diagram(s)	25	0	0.26	0.08	0.09	0.08
		Without diagrams	208	0	0.23	0.06	0.06	0.05
	0.61-0.80	With diagram(s)	17	0	0.20	0.04	0.06	0.05
		Without diagrams	139	0	0.11	0.03	0.04	0.03
	0.81-1.00	With diagram(s)	19	0	0.07	0.03	0.03	0.02
		Without diagrams	53	0	0.06	0.02	0.02	0.02
Higher	0.00-0.20	With diagram(s)	2	0.03	0.09	0.06	0.06	0.04
		Without diagrams	67	0	0.17	0.03	0.04	0.04
	0.21-0.40	With diagram(s)	18	0	0.16	0.04	0.04	0.04
		Without diagrams	164	0	0.15	0.02	0.03	0.03
	0.41-0.60	With diagram(s)	20	0	0.15	0.02	0.04	0.04
		Without diagrams	183	0	0.11	0.01	0.02	0.02
	0.61-0.80	With diagram(s)	17	0	0.13	0.01	0.02	0.04
		Without diagrams	197	0	0.09	0.01	0.01	0.02
	0.81-1.00	With diagram(s)	19	0	0.05	0	0.01	0.01
		Without diagrams	107	0	0.04	0	0	0.01

Discussion and conclusion

This research examined the accessibility of GCSE Science items that require students to create a visual or augment a partially provided one through analysing patterns of item omit rates.

Analyses of omit rates for items with and without diagrams by tier, subject, item position, item maximum mark and item facility value have shown that there was very little to no evidence that average omit rates were higher for items with diagrams compared to those without diagrams. Therefore, this research found no indication that items with diagrams in GCSE Science had potential accessibility issues.

Regardless of the item type (i.e., with or without diagrams), omit rates were overall higher for the foundation tier papers than for the higher tier papers. Furthermore, analysis of the disaggregated data by attainment group showed that omit rates were higher in the lower attainment groups and decreased in the higher attainment groups. In terms of patterns of omit rates by item position within a paper, there appeared to be a trend of increasing omit rates as a paper progressed; omit rates tended to be higher for items towards the end of the paper. These findings on omit rates by tier and attainment group, and omit rates by item position, taken together, support results from previous studies (Clemens et al., 2015; Walland, 2024), which suggested that student ability plays a role in the rate of item omission. Lower attaining students tended to omit more items and particularly items towards the end of the paper, most likely because they ran out of time to attempt these items. While this finding provides valuable insights into the likely cause of item omission, this does not indicate accessibility issues.

In terms of omit rates by subject, chemistry items with diagrams in the foundation tier papers were found to have slightly higher omit rates than those without diagrams. On the contrary, biology items with diagrams tended to have lower omit rates than those without. This finding provides an indication that subject area could also contribute to variability in omit rates in addition to or instead of item type (i.e., with or without diagrams). There could be various reasons for this, such as intrinsic differences in the kinds of visuals that learners are asked to create or augment in different subjects.

Items with diagrams that had a maximum mark of 1 in the foundation tier papers were found to have higher omit rates than those without, while items with diagrams that had a maximum mark of 3 and 4 or above had lower omit rates than those without. It could be speculated that candidates were more likely to attempt items with diagrams that had higher tariffs given the opportunity cost for not attempting them at all, while it was less of a loss for not attempting items with diagrams with lower tariffs. However, further evidence would be needed to confirm this hypothesis.

Although it does not directly concern accessibility, our finding relating to the overall patterns of item omission across gender groups is noteworthy. Male candidates in the foundation tier papers had a higher tendency to omit items than their female peers did, and this was true for both items with and without diagrams. This finding corroborates the results from Matters and Burnett (1999) indicating that the rate of omission was higher among male candidates than female candidates for constructed-response items.

It should be noted that this study was conducted using a limited dataset, based only on one exam series and a relatively small number of items involving diagram creation or augmentation. Despite this limitation, this research has demonstrated the potential of analysing omit rates to provide initial indications of item accessibility. However, in such analysis, other factors that can also influence omit rates, as discussed in this article, need to be kept in mind. A follow-up study involving a larger dataset would allow more fine-grained analyses of how the interactions between variables could potentially contribute to patterns of omit rates. Additionally, a larger dataset with more items with diagrams would enable further distinction between items involving diagram creation and those involving diagram augmentation. There could potentially be differences between items that require students to draw a diagram and items that require them to augment a partially provided one in terms of the nature and level of response strategy demand (see Pollitt et al., 2007). Items that require students to augment a partially provided diagram potentially pose less risk of a student being entirely unable to make an attempt as at least some of the diagram is already provided. It could be speculated that such differences may have implications for accessibility. In this study we could not further distinguish these two item types (create versus augment) as the numbers of items would have been even smaller. A larger dataset could allow examination of potential implications of these two item types for accessibility. Future studies should also consider gathering insights from students after they take an exam about why they leave out certain questions, to explore the contribution of different variables to omit rates and help establish the contribution of accessibility to patterns of omission.

Acknowledgement

The author would like to thank Jeff Heath for carrying out the coding of the items included in the study reported in this article.

References

Ainsworth, S., Prain, V., & Tytler, R. (2011). Drawing to learn in science. *Science*, *333*(6046), 1096–1097.

Beauchamp, D., & Constantinou, F. (2020). Using corpus linguistics tools to identify instances of low linguistic accessibility in tests. *Research Matters: A Cambridge Assessment publication, 29,* 10–16.

Beddow, P. A. (2012). Accessibility theory for enhancing the validity of test results for students with special needs. *International Journal of Disability, Development and Education, 59*(1), 97–111.

Beddow, P. A., Elliott, S. N., & Kettler, R. J. (2013). Test accessibility: Item reviews and lessons learned from four state assessments. *Education Research International, 2013*(1), 952704.

Ben-Shakhar, G., & Sinai, Y. (1991). Gender differences in multiple-choice tests: The role of differential guessing tendencies. *Journal of Educational Measurement, 28*(1), 23–35.

Chang, H. Y., Lin, T. J., Lee, M. H., Lee, S. W. Y., Lin, T. C., Tan, A. L., & Tsai, C. C. (2020). A systematic review of trends and findings in research employing drawing assessment in science education. *Studies in Science Education, 56*(1), 77–110.

Clemens, N. H., Davis, J. L., Simmons, L. E., Oslund, E. L., & Simmons, D. C. (2015). Interpreting secondary students' performance on a timed, multiple-choice reading comprehension assessment: The prevalence and impact of non-attempted items. *Journal of Psychoeducational Assessment, 33*(2), 154–165.

Crisp, V., & Macinska, S. (2020). Accessibility in GCSE Science exams - Students' perspectives. *Research Matters: A Cambridge Assessment publication, 29,* 2–10.

Department for Education. (2015). *Biology, Chemistry and Physics GCSE subject content*.

LaDue, N. D., Libarkin, J. C., & Thomas, S. R. (2015). Visual representations on high school biology, chemistry, earth science, and physics assessments. *Journal of Science Education and Technology, 24,* 818–834.

Matters, G., & Burnett, P. C. (1999). Multiple-choice versus short-response items: Differences in omit behaviour. *Australian Journal of Education, 43*(2), 117–128.

OCR. (2018a). *GCSE (9–1) Gateway Science: Exploring our question papers.*

OCR. (2018b). *GCSE (9–1) Twenty First Century Science: Exploring our question papers*.

Ofqual. (2022). *Guidance on designing and developing accessible assessments: Consultation decisions.*

Papanastasiou, E. C. (2020). Can non-responses speak louder than words? Examining patterns of item non-response in TIMSS 2015. *International Journal of Quantitative Research in Education, 5*(2), 157–172.

Pollitt, A., Ahmed, A., & Crisp, V. (2007). The demands of examination syllabuses and question papers. In P. Newton, J-A. Baird, H. Goldstein, H. Patrick, & P. Tymms (Eds.), *Techniques for monitoring the comparability of examination standards* (pp. 166–206). Qualifications and Curriculum Authority.

Posit team. (2023). *RStudio: Integrated Development Environment for R* (Version 2023.12.0.369) [Computer software]. Posit Software, PBC. http://www.posit.co/

Prain, V., & Tytler, R. (2012). Learning through constructing representations in science: a framework of representational construction affordances. *International Journal of Science Education, 34*(17), 2751–2773.

Sarac, M., & Loken, E. (2023). Examining patterns of omitted responses in a large-scale English language proficiency test. *International Journal of Testing, 23*(1), 56–72.

Trumbo, J. (1999). Visual literacy and science communication. *Science Communication, 20*(4), 409–425.

Tytler, R., Prain, V., Aranda, G., Ferguson, J., & Gorur, R. (2020). Drawing to reason and learn in science. *Journal of Research in Science Teaching, 57*(2), 209–231.

Tytler, R., Prain, V., & Hubber, P. (2018). Representation construction as a core science disciplinary literacy. In K. S. Tang & K. Danielsson (Eds.), *Global developments in literacy research for science education* (pp. 301–317). Springer.

Unsworth, L., & Herrington, M. (2023). Visualization type and frequency in final year high school science examinations. *Research in Science Education, 53*, 707–725.

von Schrader, S., & Ansley, T. (2006). Sex differences in the tendency to omit items on multiple-choice tests: 1980–2000. *Applied Measurement in Education, 19*(1), 41–65.

Walland, E. (2024). Exploring speededness in pre-reform GCSEs (2009 to 2016). *Research Matters: A Cambridge University Press & Assessment publication, 37*, 57–73.

Wang, C., & Wei, B. (2024). Analysis of visual-based physics questions of the senior high school entrance examination in China. *Physical Review Physics Education Research, 20*(1), 010112.