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/

Learners' annotations and written markings when taking a digital multiple-choice test: What support is needed?

Victoria Crisp, Sylvia Vitello, Abdullah Ali Khan, Heather Mahy and Sarah Hughes

To cite this article: Crisp, V., Vitello, S., Khan, A. A., Mahy, H., & Hughes, S. (2025). Learners' annotations and written markings when taking a digital multiple-choice test: What support is needed? *Research Matters: A Cambridge University Press & Assessment publication, 39*, 90–109. https://doi.org/10.17863/CAM.116171

To link this article: https://www.cambridgeassessment.org.uk/Images/research-matters-39-learners-annotations-and-written-markings-when-taking-a-digital-multiple-choice-test.pdf

Abstract:

This research set out to enhance our understanding of the exam techniques and types of written annotations or markings that learners may wish to use to support their thinking when taking digital multiple-choice exams. Additionally, we aimed to further explore issues around the factors that contribute to learners writing less rough work and markings on scrap paper during a digital test than they write on paper-based tests, as observed in prior research.

In this research, 52 learners attempted a digital economics test with access to either scrap paper or a print of the test. Some learners were observed in order to capture their interactions with the paper materials, all learners completed a questionnaire, and most learners were interviewed.

The evidence collected provides insights regarding the types of annotations and written markings learners wished to use. Considerable variation was found in whether, and the extent to which, learners used paper materials. Scrap paper worked fairly well for some types of annotations or written markings, but not for others. The findings are informing additional developments to testing platform functionality.

Learners' annotations and written markings when taking a digital multiple-choice test: What support is needed?

Victoria Crisp (Research Division), **Sylvia Vitello** (Research Division), **Abdullah Ali Khan** (Education Digital Products and Services), **Heather Mahy** (Education Digital Products and Services) and **Sarah Hughes** (Education Futures Directorate)

Introduction

While digital tests and exams are prevalent in many contexts, so far they are uncommon for general qualifications in England and for general qualifications available internationally that are based on the English assessment model. However, moves towards providing digital exams for appropriate general qualification contexts are now progressing at pace. Digital exams offer a variety of potential advantages over paper-based exams, for example: providing a better match to how learners conduct their school work in some subjects; assessing skills that are difficult to assess authentically in a paper-based exam (e.g., computer programming); providing customisable accessibility features that candidates can adjust to their needs (e.g., screen background colour); allowing each candidate to individually control the playback of an audio or video stimulus; and making it easier for candidates to edit their responses. Nonetheless, care is needed to ensure quality and fairness as digital exams are introduced. One of many factors that needs to be considered is ensuring that candidates are not hindered in how they work during the exam and that the digital testing platform, its functionality, and any accompanying support materials (e.g., scrap paper) allow candidates to use their relevant knowledge, understanding and skills to select or produce their answers. This relates to validity since it contributes to ensuring that candidates' results reflect relevant constructs and can be interpreted and used in the intended ways (Messick, 1989).

As well as ensuring validity, comparability also needs to be considered. In contexts where parallel digital and paper-based exams are offered as alternatives, careful thought needs to be given to the intentions for comparability between modes (see Shaw, Crisp & Hughes, 2020, for a framework to support thinking about the different kinds of claims that might be made regarding comparability between assessments). One important aspect of this is the extent to which cognitive

processes (as supported by tools and materials) differ between modes and whether any differences are appropriate given the comparability claims that are made.

These considerations around validity and comparability have led to discussion within Cambridge University Press & Assessment about the annotations that candidates make on their exam papers in paper-based exams, the functions that these serve in terms of supporting candidates' question-answering processes and how such functions can be appropriately supported when candidates take digital exams. As a starting point, Williamson (2025, this issue) analysed candidates' annotations in paper-based GCSE Maths and Science exam scripts. She found that annotations occurred quite frequently (overall rate of 40 per cent across all questions considered) and that rates of annotation varied for different questions (from 8 to 93 per cent). The types of annotations observed to have been made by candidates on their paper-based exams included highlighting key information, crossing or ticking response options in multiple-choice questions (MCQs), annotating the question with related facts or rules, annotating a graph or figure, and showing working out.

Several other studies have explored how learners use scrap paper to support their exam techniques when taking digital tests, often with either the same learners or parallel groups of learners attempting digital and paper-based versions of the same or similar tests in order that comparisons can be made. One finding is that some learners transfer material (e.g., diagrams) from screen to scrap paper when taking a digital test so that they can annotate, and that this transfer can sometimes lead to errors (Johnson & Green, 2006; Hughes et al., 2011). Another important finding is that learners tend to write less rough work on scrap paper during digital tests than on the test paper during a paper-based test (Johnson & Green, 2006; Hughes et al., 2011; Pengelley et al., 2023; Nastuta & Liu, 2023). It has been theorised (using cognitive load theory) that this may be a result of an additional cognitive load cost that may be incurred when switching attention between modes (Pengelley et al., 2023).

Another possible contributor to reduced written working and markings on scrap paper during digital tests is simply that certain written actions are not possible, or not as easy or natural, on scrap paper compared to on a paper-based test. For example, to annotate a diagram provided in a question, as exemplified in Hughes et al. (2011), learners taking a digital test would have to copy the diagram to scrap paper before being able to add annotations. Settlage and Wollscheid (2024) explored how being able to write on a paper test might contribute to mode effects. They allowed one cohort of learners taking three paper-based multiple-choice tests to write on their test papers while another cohort (taking the same paper-based tests) were only allowed to write on scrap paper. After controlling for differences in ability, the findings showed that those who were allowed to write on their test paper performed significantly better on two out of three tests, and overall. Performance improved by 3.5 per cent overall. Settlage and Wollscheid (2024) do not report the volume of written work from each cohort, except to note that they were surprised that over 90 per cent of learners who were instructed not to write on the test papers wrote nothing on their scrap paper. Their findings

imply that mode switching is not the only factor affecting the extent to which learners make annotations and written markings on scrap paper during digital tests.

The current research was specifically motivated by Cambridge International Education's plans to introduce digital versions of IGCSE multiple-choice exams in economics, accounting and the sciences running in parallel to existing paper-based exams. As well as potential advantages for the efficiency of processing and marking responses, for candidates it removes the current need to record their responses on a machine-readable form separate to the exam paper. To support this development, this research set out to enhance our understanding of the exam techniques and types of written annotations or markings that learners may wish to be able to use to support their thinking when taking digital multiple-choice exams. It was hoped that the findings would provide insights to inform any necessary additional developments to testing platform functionality and inform decisions about the need for any accompanying materials (e.g., scrap paper). Additionally, the research aimed to further explore issues around the factors that contribute to learners writing less rough work and markings on scrap paper during a digital test than they write on paper-based tests. To explore these themes, we asked learners to take a digital multiple-choice test (based on IGCSE Economics questions) while having access to either scrap paper or a print of the test. The inclusion of a test print was also interesting from the perspective that this could potentially be an option that testing organisations could consider providing to learners early in the introduction of digital exams, or to learners with certain kinds of learning needs. That said, a test print may not be an elegant solution to supporting learners' cognitive processes during a digital exam, given it involves duplication of material.

Since written (or sketched) work on scrap paper cannot technically be considered annotation (unless a learner reproduces part of the question or stimulus first), in this article we refer to both "annotation" and "written markings" to mean any writing (or sketching) on the test print or on scrap paper.

Method

Participants

The participants were 52 learners in three schools in England, aged around 17 years old, who were studying A Level Economics. Unfortunately, it was not possible to involve learners studying IGCSE Economics due to practicalities around timing and access. As a result, the test questions are likely to have been slightly easy for the learners, which should be kept in mind when interpreting the findings.

Materials

Digital test

A digital multiple-choice economics test was prepared using 15 questions from a past IGCSE exam paper (designed for 16-year-olds). The questions included a range of common design features, such as a stimulus diagram, a stimulus table, calculation, text-only response options, and tabulated response options.

In addition, a short digital test containing three other multiple-choice economics questions was prepared for demonstration purposes.

Test print

A PDF of the test (downloaded from the testing platform) was printed, double-sided on A4 paper.

Scrap paper

Scrap paper booklets were created to include two A4 sheets of lined paper followed by two A4 sheets of plain paper.

Procedure

The procedure was as follows:

- **Introduction** — The research and what would be involved in participation was described to learners, and informed consent was gained.
- **Platform demonstration** — The testing platform was demonstrated to learners using the demonstration test. Learners were shown navigation and other available functionality (i.e., a notes tool, calculator, question flagging tool).
- **Paper materials assigned** — Each learner was assigned to one of two conditions by being given either a print of the test or a scrap paper booklet, at random. It was emphasised to learners that they could use this paper as much or as little as they liked.
- **Test** — Learners were given up to 25 minutes to attempt the digital test. They used their own devices (laptops; tablets with or without keyboards).
- **Observations** — During the test, some learners were observed by a researcher, with notes on each learner's interactions with screen and paper captured in an observation sheet.
- **Questionnaire** — After the test, all learners completed an online questionnaire. This asked about their views and experiences of the digital test, their exam techniques and ways of working, and their use of the paper materials.
- **Interviews** — After the test and questionnaire, most learners were interviewed. Each interview usually involved a pair of learners where each learner had been given a different type of paper material. The interviews were semi-structured and explored learners' views and experiences in more depth.

The numbers of learners who participated in each element of the research are shown in Table 1. Learners were assigned random IDs, in the form "L01".

Table 1: Number of learners taking part in each research task

Research task	Number of learners		
	Test print	Scrap paper	Total
Digital test	27	25	52
Observation during test[1]	15	9	24
Post-test questionnaire	27	25	52
Post-test interview	22	23	45

1 The imbalance in the numbers of learners in each condition who were observed arose because of practicalities in the classrooms during data collection sessions.

Ethical considerations

The research protocol was reviewed by Cambridge University Press & Assessment's Research Ethics Committee and received a favourable outcome. Only learners who gave informed consent participated in the research.

Analysis

Initial analysis of observation notes involved counting the number of observed learners who interacted with the paper materials in any way. Then, further analyses explored how frequently learners used the paper when attempting questions, how much learners switched between modes, and which mode they spent more time on. Many of the observed learners carried out a main round of attempting the test questions, followed by a round of checking their responses. The analyses focused on each learner's main round of attempting the questions.

The paper materials given to learners for use during the test were inspected for whether there was any written work, drawing or annotations and, if so, how much space was used. Additionally, the kinds of annotations and written markings made by learners were categorised.

To analyse the questionnaire data, frequencies and percentages of each response were obtained. The interview transcripts were analysed using thematic analysis.

Findings

The findings are organised by overarching theme, with evidence from the different data sources included as relevant. We begin with evidence on the amount of written work conducted by learners and the possible effect of the nature of the assessment on this. We then look at the types of annotations and written markings that learners reported using in paper-based tests, and that they used during the digital test taken for the research. Next, we consider learners' views on the roles served by annotations and written markings, on having access to paper during the digital test, and on the functions that paper materials can support. Finally, we cover how learners interacted with the paper materials, the impact of switching attention between modes, and the impact of mode on the amount of mental working.

Amount of written work

In total, 25 out of the 52 learners (48 per cent) made some written markings on the paper materials they were given. This same proportion was found in both paper conditions (13 out of 27 learners in the print condition and 12 out of 25 learners in the scrap condition wrote on the paper).

Learners' paper materials were inspected for markings relating to each test question. This revealed wide variation between learners in terms of how many questions they made annotations for (Figure 1). A small number of learners, most of whom received the test print, made annotations for all, or almost all, questions. Other learners made annotations for less than half of the questions. Among these learners, those with scrap paper tended to annotate for fewer questions than those with the test print.

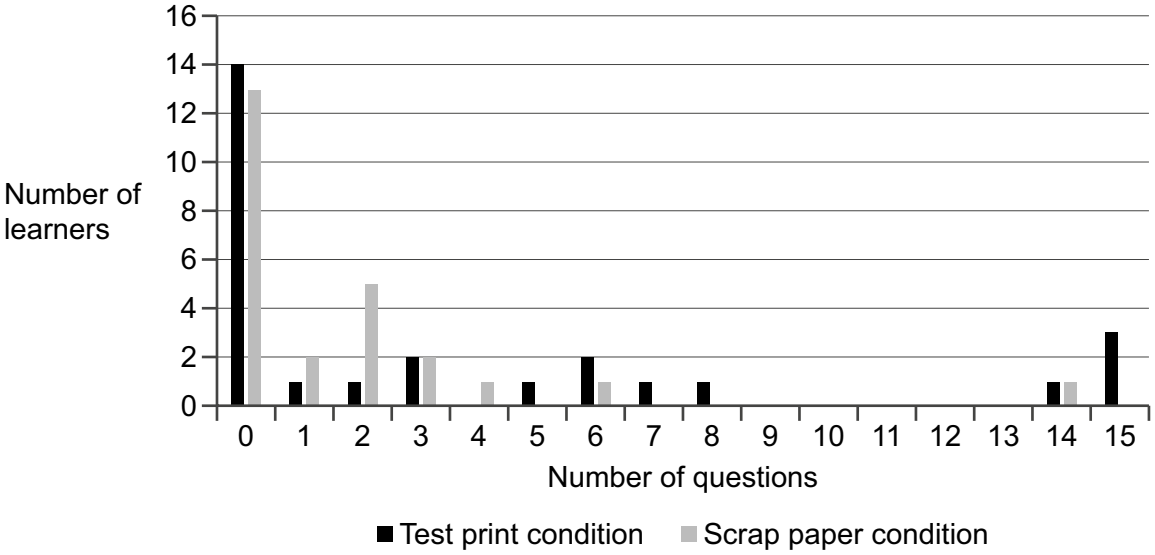

Figure 1: The number of questions for which learners made any written markings on paper

The analysis also considered how much physical space learners used on the paper materials. In the print condition, of the 13 learners who wrote on the paper, eight used moderate amounts of space in or around some questions, two used minimal amounts of space in or around one or two questions, and one used large amounts of space in or around most questions. The remaining two learners used blank space on the first page of the test print as if it were scrap paper (i.e., their written markings were not on or next to the relevant questions).

In the scrap paper condition, 10 of the 12 learners who wrote on paper used one side of the lined paper in the scrap paper booklets provided to them. Five of these learners used less than a third of a side, while the other five used more than 50 per cent of one side. Two learners wrote on two or three sides of lined paper.

None of the learners used the plain paper that was provided after the lined paper.[2]

Effect of assessment context on paper use

Some comments made during interviews suggested that learners might have used the paper materials differently if the test had been different in some respect. Learners' points are summarised in Table 2. Comments suggest that the paper materials may have been used more if the questions had been more difficult, involved more complex diagrams or calculations, or if the test had higher stakes.

2 From this evidence, we cannot tell whether learners with scrap paper would have used the plain paper rather than the lined paper if the plain paper sheets had been arranged first in the scrap paper booklets. The questionnaire asked the learners who had received scrap paper for their views on the type of paper they would prefer to have when taking a digital test. Most preferred lined paper (11 learners) or did not mind what kind of paper they received (seven learners). A few learners preferred plain paper (three learners) or reported that they would like both lined and plain paper (three learners). (One learner gave no response.)

Table 2: Learners' interview comments regarding the effect of the nature of the test or questions on their use of paper materials (summarised comments)

Test print	Scrap paper
• Did not need the print for this test but useful for other tests. • Likely to use the test print more for more difficult questions. • Might have looked at the test print if there had been a more complicated graph to analyse. • Did not need the print for this test but it would be useful for a big graph or a table with lots of data. • Might draw diagrams on the print for A Level MCQs. • For A Level long-answer questions they might use the test print to annotate the question or draw the structure of their point of view or evaluation system.	• More likely to use scrap paper for more complex or more difficult questions. • Might use scrap paper if planning a longer answer/essay. • Might use scrap paper for more difficult calculation questions. • Scrap paper is useful when they need to draw a diagram or graph. • Might have drawn diagrams on the scrap paper if the test had been high stakes. • Considered using the scrap paper for a question involving an equation or calculation, but then found it was easy to do without any written work.

Types of annotations and written markings

In the questionnaire, learners were asked which exam techniques, from a list provided, they usually use when taking paper-based tests (Table 3). Ruling out incorrect responses was most commonly selected, followed by sketching a diagram, graph or table, underlining or circling key words, annotating an existing diagram, graph or table, showing working out, and keeping track of questions to revisit later.

Table 3: Usual exam techniques used in paper-based tests (closed response)

What exam techniques do you usually use in paper tests, including when answering multiple-choice questions? Select all that you use.	**Number and percentage of learners who selected each option**	
I underline/circle key words	34	65.4%
I write down facts, formulae or theories related to the question	29	55.8%
I make notes on ideas	18	34.6%
I show my working out	32	61.5%
I rule out answers (A, B, C, D) as soon as I know they're wrong	46	88.5%
I annotate the graphs, diagrams or tables that are in the test	33	63.5%
I sketch my own graph, diagram or table	34	65.4%
I keep track of questions that I want to come back to	32	61.5%
I don't do anything in particular	1	1.9%
Other (please specify)	0	0.0%

The paper materials used by learners during the research provide insights into the kinds of written markings made by learners while attempting the digital test. Categories were developed to capture the types of markings that were apparent. Figure 2 shows the numbers of learners who made markings fitting each category.

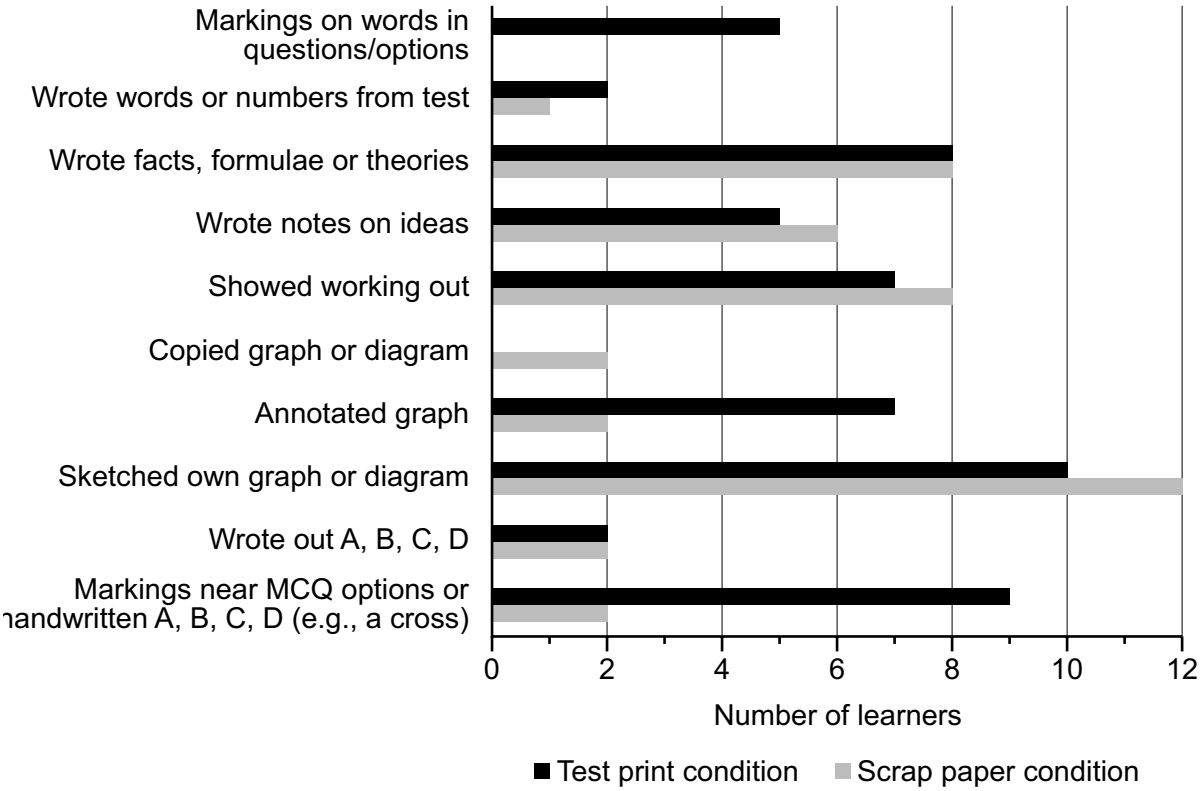

Figure 2: The number of learners in both paper conditions who made each type of annotation or marking on their paper. The categories are ordered thematically.

In both conditions it was relatively common for learners to sketch their own graph or diagram, show their working out, and to write down facts, formulae or theories. Other types of annotation or written markings were notably more common among learners in the test print condition than in the scrap paper condition. The most marked difference was found for markings on or near multiple-choice response options, presumably made to record ruling in or out a response option. Only two learners with scrap paper made any markings related to selecting response options (e.g., writing out the letters "A", "B", "C", "D" and then circling a letter or using a line to strike through some letters). Annotating a graph was also more common in the test print than on scrap paper, perhaps unsurprisingly since it was only possible on scrap paper if learners copied the graph first. Additionally, learners with the test print were more likely to identify key information (e.g., words, numbers) in the question or response options (by circling or underlining) than learners in the scrap paper condition (by writing down words).

For each category in Figure 2, the data showed a range of specific, and sometimes idiosyncratic, annotations or markings. To provide an illustration of this variation, Figure 3 and Figure 4 show some examples relating to response options, as made on the test print and on scrap paper.[3] As can be seen, learners used ticks, crosses, circles and strikethrough lines as part of their process of ruling in or out options.

3 See Williamson (2025, this issue) for illustrations of other types of annotation in paper-based exams.

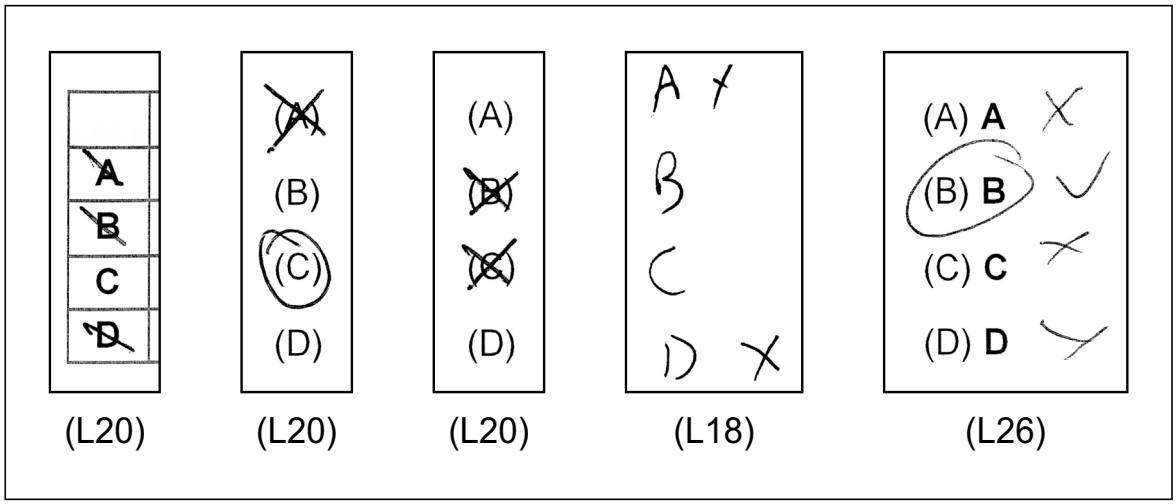

Figure 3: Examples of annotations and written markings relating to multiple-choice response options made in the test print[4]

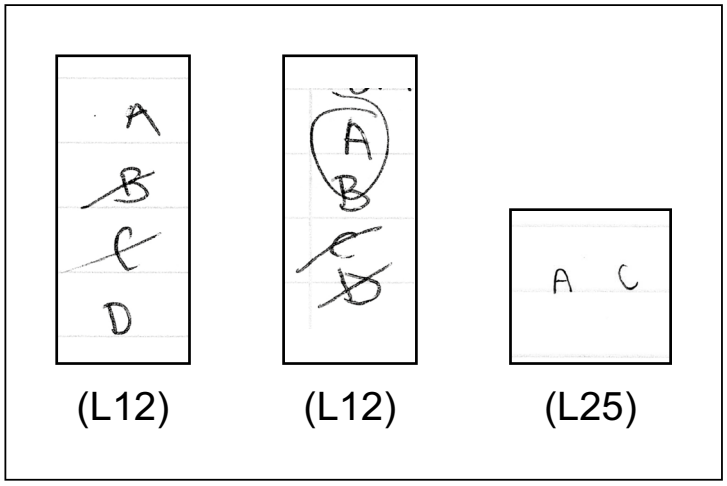

Figure 4: Examples of annotations and written markings relating to multiple-choice response options made on scrap paper

To summarise the variety of different kinds of written markings that were observed, Table 4 groups the categories from Figure 2 into five broader themes and lists illustrative examples. There was variation in the number of questions for which an individual learner used a particular type of annotation or written marking across questions. In some cases frequency of use was influenced by whether an annotation or marking type was appropriate to all or only to some questions (e.g., showing working out was only relevant to questions that involved calculation or a formula).

4 Note that learner L18 made their written markings on the first page of the test print.

Table 4: Annotation and marking types organised by theme, with exemplification

Theme	Examples
Extracting information from the question	• Underlining and circling key words. • Writing down key words from question context. • Writing down numbers required for calculations. • Copying information (e.g., percentages) from a table (without reproducing the table format).
Writing notes on subject content and ideas	• Writing down recalled economics formulae. • Writing down economic theory acronyms. • Writing down economics study mnemonics. • Using symbols to present information (e.g., up/down arrows). • Writing thoughts about the question or answer in short phrases or sentences.
Showing working out	• Arithmetic calculation steps. • Working out an economics formula numerically (e.g., inputting numbers). • Working out an economics formula conceptually (e.g., writing comments alongside it).
Graphs and diagrams	• Adding points (e.g., dots or circles) to a graph in the test question. • Adding arrows to a graph in the test question. • Shading areas of a graph in the test question. • Sketching one or more versions of a new graph. • Sketching a number line.
Response options	• Writing down letters "A" "B" "C" "D". • Putting a cross on (or next to) response options, presumably to rule them out. • Striking through options with a line, presumably to rule them out. • Circling one response option, presumably to indicate the correct response. • Putting a star next to one option, presumably to indicate the correct response. • Putting markings in tables next to response options (either to rule out options or mark-up plausible options). • Putting a "?" next to options.

Roles served by annotations and written markings

The interviews provide evidence regarding learners' views on the roles played by written markings in their question answering. One key theme was that annotations or written markings could support learners' thinking or working out and allow them to see their whole chain of reasoning. For example:

> "I mean for me I find it easier to kind of just write down my chain of thinking and reasoning and just having that in front of me ... when it is in front of me, I can visualise it, it is a bit easier to kind of translate that into how I would work out the answer. So I feel that, for me, it's quite important, especially when the questions are a bit more complex." (L25, scrap paper)
>
> "I'm the kind of person that needs to write down all my thoughts." (L46, test print)

Two learners (interviewed together) felt that making notes could help unpack the meaning of key terms:

> "I would find a word and then I had to figure out like the key term. 'Oh, 'inelastic' blah blah blah' and like figure that out and write it down and then follow with my train of thought." (L12, scrap paper)

For another learner, their usual practice of writing notes helped identify which of their ideas were relevant:

> "I had nothing to scribble down any ideas. Because, even if I don't necessarily use what I scribble down, it kind of helps me filter through information to see what is relevant to the question." (L24, test print)

A common theme was that being able to annotate a stimulus graph or diagram, or to sketch a stimulus graph or diagram, supported their thinking, for example by allowing them to visualise concepts:

> "For one of the questions, I did use a really simple diagram that I drew to sort of just help me visualise what I think was the most prominent sort of choice." (L23, scrap paper)

> "Normally if I have a question with the graph, I'll try to annotate the graph ... it was a bit difficult to try and understand like, understand my answer without being able to draw on it." (L38, scrap paper)

Some learners commented on how annotations played a role in ruling in or out response options when answering multiple-choice questions:

> "I can't cross out the wrong answer on the computer and I also can't show my workings on the computer, so it's going to make my chain of analysis like a bit messy because I have to do everything in my brain." (L45, test print)

> "I feel better about myself if I know that I've only got two options left at one point ... when I'm able to cross my final one, I can like really focus on just my last option to make sure that I've got it right." (L24, test print)

Views on having paper materials during the digital test

In the questionnaire, learners who reported that they had not used the paper materials were asked whether they liked having the paper or would be happy not to have it. Eight of the 14 learners who did not use their test print, and 12 of the 13 learners who did not use their scrap paper, recorded that they liked having it. The remaining seven learners who had not used their scrap paper or test print reported that they would be happy not to have the materials they received.

Additional evidence comes from the interviews, during which learners were asked how they felt about having the paper materials they received. For each paper type, some learners reported that they liked having the paper materials and having them was comforting or reduced stress:

> "A sense of comfort because I know, if I ever need it, it's there for me to use." (L36, scrap paper)

> "I like having them, it was comforting." (L19, test print)

One learner felt that not having scrap paper would have affected their performance (presumably in a negative way):

> "I think, if I didn't have it, it would have – it would have affected my, like, answers." (L52, scrap paper)

Small numbers of learners in each condition felt that they should use the paper because they had been given it:

> "I kind of felt like I needed to use it, so I did. ... Because it's like, I don't know, it's just like there." (L27, test print)

Views on the functions that paper materials can serve during a digital test

During interviews, learners commented on the perceived benefits of the two types of paper materials, which often related to their understandings of how paper materials can support different types of written working and annotation. Table 5 summarises these comments. Scrap paper reportedly served various functions such as allowing learners to sketch diagrams and write down calculations or ideas, and supporting their thinking. For the test print, learners showed an awareness of the additional opportunities that this offered, such as annotating the text or a diagram, avoiding the need to re-draw diagrams, and making annotations near the relevant question.

Table 5: Learners' interview comments regarding benefits of the paper materials (summarised comments with some quotations for exemplification)

Test print	Scrap paper
• It can be annotated (e.g., underline key words, draw on diagrams, cross out options). • Aids quick annotation. • Annotations/notes can be made near the question so less confusing than scrap paper – you know which notes relate to which question. • More useful than digital notes tool which is separate from the text. • Reduces the need to switch between screen and paper, e.g., "I liked having it printed out because I could read it, and I didn't have to be going like back and forth between reading the question online" (L19, test print). • Avoids needing to re-draw diagrams, which would take up time. • Allows use of a pen to point at or focus attention on parts of the question. • Appropriate amount of space around questions to work. • Useful for checking/reviewing answers, e.g., "just an easier way to get my thoughts in check" (L26, test print). • Useful if there are multiple steps or if question involves numbers. • Easier to process the questions if you can write things down rather than doing it in your head. • Able to write down thoughts and do maths on paper.	• Useful for drawing diagrams/graphs (e.g., helps to visualise the answer). • Useful for calculations (e.g., writing down a multiple-step calculation). • Easier than doing working in head as they can see the work in front of them, which supports their thinking. • Can record train of thought / write down working out as they go. • Can write anywhere and get ideas down quickly. • Provides space for rough work. • More familiar with using scrap paper. • Would rather write on paper than type (e.g., using digital notes tool).

How learners interacted with paper materials during the digital test

The observations of individual learners allowed a more detailed exploration of the ways that learners interacted with the paper materials for each test question. In total 13 of the observed learners interacted with their paper materials in some way (nine of the 15 observed learners who had the test print, and four of the nine observed learners who had scrap paper). This includes learners who only looked at or touched the paper, as well as learners who wrote or made markings on the paper.

For the 13 observed learners who interacted with the paper materials in some way, the observation notes provide insights regarding:

- Level of interaction with paper – whether the learner interacted with the paper including annotating or making markings, interacted with the paper but only by looking at or touching it, or did not interact with the paper.
- Integration of the two modes – whether the learner started reading the question on screen or on paper (as an indication of whether their processing of the question was led by the screen or paper mode), and how many times the learner "visited" the non-leading mode from their lead mode (no visits, one visit, or multiple visits). Only visits that involved some processing of the question were counted (i.e., visits to screen were not counted if learners only carried out administrative tasks such as clicking the platform's "next" button to advance to the next question).
- Relative time on the two modes – whether the learner seemed to have spent more time attending to the screen, more time attending to paper, or a similar amount of time on each mode. This was based on subjective judgements made by the researchers during observations.

Figure 5 shows these three aspects of interaction side by side for each relevant learner. The left panel shows considerable variation between learners in the number of questions for which they interacted with paper. When learners interacted with paper, in most cases they made some written markings.

There was also variation between learners in terms of how they used the two modes (middle panel). Only one learner used paper as their lead mode. They read all questions on their test print and did not make any visits to the screen, except to input their responses at the end of the test. The other 12 learners were all screen-led for all test questions, but they varied in how much they switched between screen and paper.

The right panel shows that learners differed in how they divided their time between modes, although there seems to be less variation than in other respects. Two learners spent more time on paper than on screen for all, or almost all, questions. Four learners spent more time on screen for all questions. The remaining learners spent more time on screen for most questions, but for some questions spent more time on paper or a similar amount of time on both modes.

The variation found between the observed learners does not seem to depend on paper condition. Learners with scrap paper have patterns that would not look out of place among the patterns found for the learners with the test print.

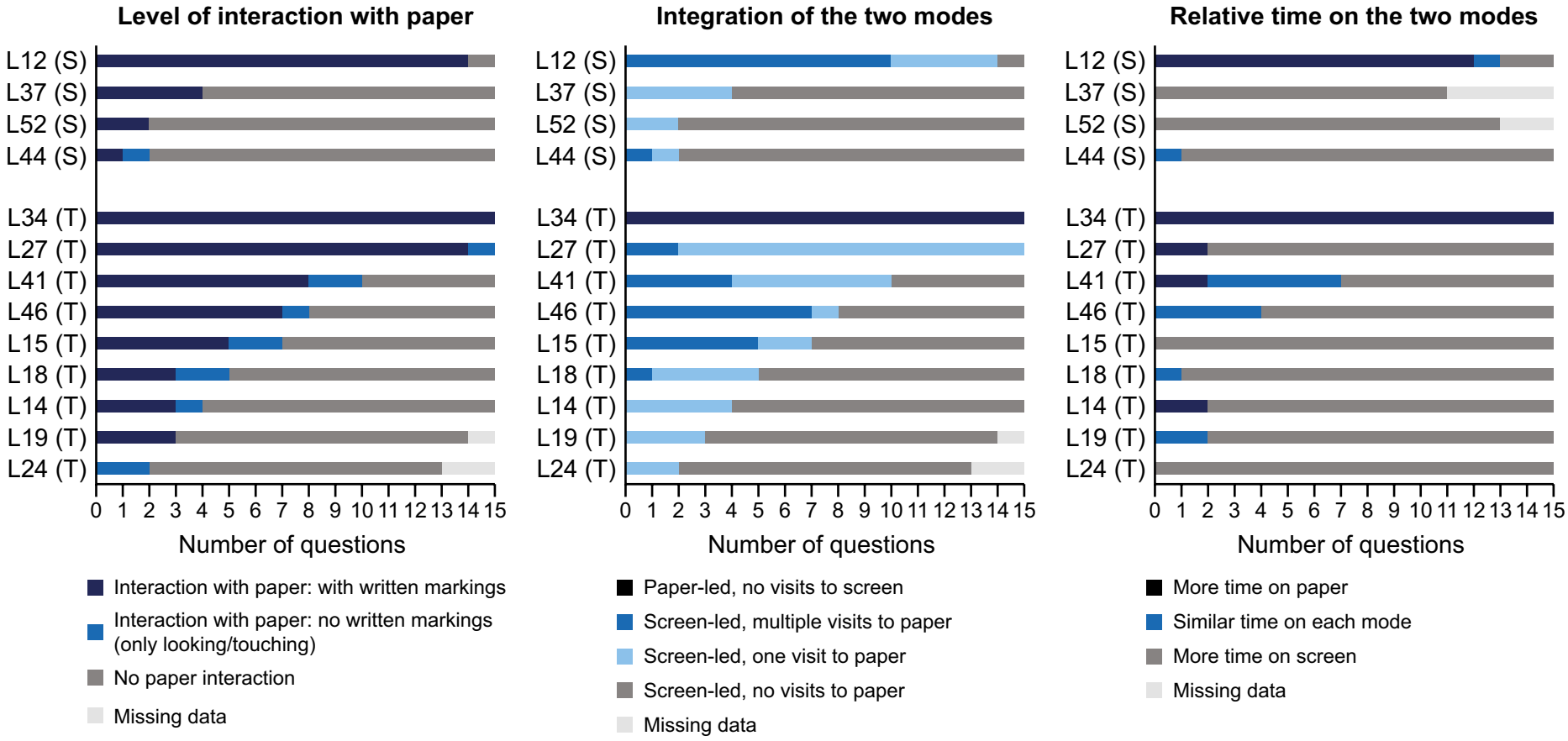

Figure 5: Profiles of interaction with paper for each of the 13 observed learners who interacted with the paper in some way.
S = Scrap paper condition; T = Test print condition.

To provide some relevant evidence about interactions with paper for all learners (i.e., not just those whom we observed), the questionnaire asked learners who received the test print whether they tended to read the test questions on screen or on paper first. Three learners stated that they usually read the questions on paper first, one responded that they read some questions on screen first and some on paper first, and 22 reported that they usually read questions on screen first.[5] This echoes the observation findings in the middle panel of Figure 5, and again suggests that using the test print as if it were a paper-based test was relatively uncommon.

Impact of switching attention between digital and paper modes

While analysis of the observations indicated the extent to which learners moved their attention between modes (Figure 5), learner comments during interviews provide insights into their experiences and views regarding mode-switching. Several comments suggested that switching modes was slightly inconvenient or confusing:

> "I had to continually switch from like scrap paper, question paper to iPad which was kind of inconvenient." (L46, test print)

> "It's a bit different having scrap paper and the test online because I have to look at the computer and look at my paper, it's just a tiny bit confusing but it's fine, yeah." (L05, scrap paper)

A related point was that switching modes appeared to act as a barrier to writing on the paper materials for some learners and that making notes or sketches would have been easier and more likely if this could be done in close proximity to the question:

> "If the test was on paper, like totally, then I'd probably write my workings out a little more." (L18, test print)

> "I didn't really do any [written work]. I might- I might have done some if it was on paper because then it's just easier I guess when the question's right there. I might have just drawn like a diagram or what numbers I use for a calculation. But it wasn't changed drastically." (L07, scrap paper)

Two learners (interviewed together) noted that switching between paper and screen interrupted their thinking:

> "If you look away, you look back, and you're, like: 'What was I thinking? Where was I looking?'" (L31, test print)

Another point about switching raised by three learners was that when answering a paper-based test they would already have a pen in their hand and that this, in itself, might increase the likelihood of annotation compared to when taking a digital test:

> "If I'm like doing it on paper, I would have a pen in my hand. It would be normal to just write it down." (L50, scrap paper)

Impact of mode on mental working

Comments from learners about doing more of the work "in their head" were common in the interviews. This often related to performing working out such as calculations, or holding ideas in working memory:

> "Usually I tend to write stuff out more, but with the computer I was more inclined to work stuff out in my head." (L18, test print)

5 The response was missing for one learner.

> "I kind of had to think more in my head and, like, remember everything that was going on." (L22, scrap paper)

One learner's comments suggested that conducting more thinking without writing anything down was partly a result of their lack of familiarity with the test being on screen:

> "I try to do it all in my head if I'm on the computer. ... I think it's because I don't have the ... thing right in front of me, where I can jot notes down or jot techniques down that I could use. I think it's just the position of the screen is slightly alien to me." (L08, test print)

Some learners noted that they did most of their working mentally during the test but that this was their usual practice for multiple-choice tests:

> "I kind of just did it in my head and I think that is what I'd normally do with multiple-choice questions." (L51, test print)

For other learners, the calculation question in the test was sufficiently easy that conducting the calculation mentally was unproblematic:

> "Even when I did the maths question I was like, 'I might write something down here', but then I was like 'I can just do it in my head, it's easy enough.'" (L21, test print)

However, a few learners worried about making mistakes, which sometimes led them to using paper to avoid this risk:

> "I also had to write down my calculations because I make a lot of mistakes when I do things in my head." (L46, test print)

Relatedly, some learners felt that it would be more difficult to keep track of their thinking or ideas when there were greater demands on working memory:

> "If you didn't [have paper] then you kind of just have to think about it in your head which can make it harder I suppose to remember what you want to say." (L51, test print)

Discussion

This research explored learners' practices in terms of annotations and written markings made on paper when taking digital multiple-choice tests. These are important considerations for ensuring that learners can effectively show their relevant knowledge, understanding and skills through the answers they select, and for ensuring comparability where paper and digital versions of an exam are offered in parallel (depending on the exact intentions regarding comparability). The research found that learners used, or would have liked to use, a range of written marking types such as making notes, writing down working out, annotating or sketching a diagram, circling or underlining words, and ruling in or out multiple-choice response options. Two key recommendations about how to support candidates' question-answering processes during digital exams arise from the research. Firstly, scrap paper should be provided to all candidates when they take a digital multiple-choice exam in economics or another subject where some questions tend to involve calculation, use of formulae or sketching visuals to support answering, as scrap paper can serve various useful functions (including

allowing candidates to use some similar strategies to those they may use during paper-based exams). Providing scrap paper during digital exams may also be appropriate for a wider range of subjects and question types. Secondly, providing easy-to-use digital functionality to support annotation is particularly important for those functions that scrap paper does not serve well (e.g., annotating a graph). While the test print better facilitated some annotation types, the percentage of learners choosing to use it was no higher than for scrap paper and, as mentioned earlier, it would not be an ideal long-term approach given the duplication of material.

Two limitations relating to our research sample should be kept in mind when interpreting the findings. The participating learners were not due to take high-stakes digital exams so had not received relevant exam preparation. Additionally, the participants were Year 12 economics learners in England and not learners who were studying for IGCSE Economics. Thus, the questions were probably relatively easy for these learners, which could have affected their behaviours. Learners tended to report that their frequency of written markings might have been higher for more difficult questions, which could suggest that IGCSE learners would have annotated more frequently than our participants. However, we cannot be sure of this, particularly as some evidence suggests that learners use scrap paper less for more difficult questions than for easier questions (Pengelley et al., 2023). Despite these limitations, the findings provide a diverse set of evidence with no indications that the types of annotations or written markings would have been different for learners preparing to take IGCSE Economics.

Williamson (2025, this issue), and questionnaire responses from the current research, show that learners use various exam techniques in paper-based exams that involve making annotations or written markings (that are not part of their responses) on the exam paper. There was considerable variation in whether and how much our research participants used paper materials when taking the digital test. This may reflect exam-taking practices that learners have been taught or their own preferred ways of working during exams that have developed over time. It might also have been affected by differences in learner ability (and, therefore, how difficult the questions were for different learners), different levels of motivation and engagement for a low-stakes experimental test, and expectations around whether they should conduct more working mentally when taking a digital test. Additionally, a lack of familiarity and preparation for using paper alongside a digital test may have affected learners differently. When learners are preparing to take high-stakes digital exams, they will have opportunities to undertake practice tests and are likely to receive guidance around the ways that they can use digital platform functionality and any accompanying paper materials for annotations and written markings. Hopefully, such preparation opportunities should help each learner use the available tools and materials in the ways that best support them.

Our findings suggest that for some types of written markings, scrap paper generally worked well. Learners could sketch their own graphs or diagrams, note down ideas, facts, formulae or memory aids, and record working for calculations. Some learners did comment, however, that it was less confusing to write next to the relevant question in a paper-based test, but fundamentally scrap paper was

able to support these functions. For some other types of written markings, scrap paper did not work well. These included: marking a key word (e.g., by underlining or circling); annotating a graph provided in the test; and making markings to support ruling in or out multiple-choice response options. Small numbers of learners made efforts to perform these functions using scrap paper (e.g., copying a graph, writing out "A", "B", "C", "D"), but this was uncommon and viewed as inconvenient.

As discussed earlier, past research suggests that learners tend to write less working on scrap paper when taking a digital test than they do on a test paper when taking a paper-based test (e.g., Hughes et al., 2011; Johnson & Green, 2006). One possible explanation is that there is an additional cognitive load cost to switching attention between screen and paper, adding to working memory demands (Pengelley et al., 2023). Some learners' comments in the current research suggested that they felt they conducted more working mentally, and that mode switching was a contributing factor. The inclusion of the test print allowed us to consider whether past findings of learners writing less on scrap paper during digital tests (compared to the amount of rough work conducted in paper-based tests) may be partially a result of certain actions not being possible (or convenient) on this type of paper material. Our findings support this as a possible explanation. Learners receiving the test print were no more likely to write on the paper support materials overall than those who received scrap paper, but certain types of annotations or written markings were more common among those with the test print.

It also appears plausible that various other factors could contribute to reduced written work during digital tests, such as learners' expectations of how they should work during a digital exam and their existing test-taking habits. For example, some may expect not to use paper at all and have little past experience of using paper while taking a digital test. Additionally, there is a physical element as well as the cognitive element of switching; for example, some learners commented that in a paper-based exam they would already have a pen in their hand.

Taken together, it seems likely that several factors contribute to fewer written annotations and markings being made on scrap paper during digital tests than are made on paper-based tests: the cognitive load cost involved in switching mode; physical factors involved in switching mode; learner expectations and habits; and the functions that scrap paper can and cannot easily facilitate. In these ways, using scrap paper during a digital exam is different in nature to using the exam paper for annotation and rough work during a paper-based exam. The current research, and other prior research, exemplifies cases where learners chose not to write on paper even though they could have. The potential for the cognitive effort of switching attention between modes and other factors to act as barriers to paper use emphasises the need for teachers and exam providers to ensure that learners have opportunities to practise using scrap paper while taking digital tests so that this becomes familiar.

Conclusion

The current findings show that scrap paper can serve some of the important functions that annotations normally serve in paper-based exams. Therefore, we would argue that scrap paper should be given to all candidates taking a digital multiple-choice exam in economics or in another subject where some questions tend to involve calculation, use of formulae or sketching visuals to support answering. The provision of scrap paper during digital exams also seems likely to be appropriate across a wider range of subjects and question types, since some of the types of written markings observed in the current research are relevant to many exam contexts. For example, writing notes on subject content and ideas may be relevant to many question types, including essays, and for some learners this may be easier on paper than on screen. Additional research could add to our understanding of the importance of scrap paper for different contexts.

It is also important that digital testing platforms include appropriate and easy-to-use functionality to support annotations, particularly for those written marking types for which scrap paper does not work well. For the IGCSE Economics questions explored in the current research, these were: identifying key words, ruling in or out multiple-choice response options, and annotating a stimulus diagram or graph. Evidence from this research is feeding into the development of Cambridge University Press & Assessment's digital exams. For example, findings are informing prototyping and user experience testing of additional annotation functionality, which will then feed into platform developments. It is possible that, in time, the provision of appropriate digital functionality and increased learner familiarity with using such functionality could reduce the range of assessment contexts for which scrap paper needs to be given to all candidates.

Acknowledgement

With many thanks to various colleagues for their input during the planning of this research, to the learners who participated and to their teachers.

References

Hughes, S., Custodio, I., Sweiry, E., & Clesham, R. (2011, November 8–10). *Beyond multiple choice: Do e-assessment and mathematics add up?* [Paper presentation]. AEA-Europe 12th Annual Conference, Belfast, Northern Ireland, UK.

Johnson, M., & Green, S. (2006). On-line mathematics assessment: The impact of mode on performance and question answering strategies. *The Journal of Technology, Learning, and Assessment, 4*(5).

Messick, S. (1989). Validity. In R. Linn (Ed.), *Educational Measurement* (pp. 13–103). Macmillan.

Nastuta, S., & Liu, L. (2023, November 1–4). *TIMSS 2019 equivalence study: A quantitative approach to explore assessment mode effects on mathematics performance in England* [Paper presentation]. AEA-Europe 24th Annual Conference, Malta.

Pengelley, J., Whipp, P. R., & Rovis-Hermann, N. (2023). A testing load: Investigating test mode effects on test score, cognitive load and scratch paper use with secondary school students. *Educational Psychology Review, 35*(3), 67.

Settlage, D. M., & Wollscheid, J. R. (2024). Deconstructing the testing mode effect: Analyzing the difference between writing and no writing on the test. *Journal of the Scholarship of Teaching and Learning, 24*(2), 79–89.

Shaw, S. D., Crisp, V., & Hughes, S. (2020). A framework for describing comparability between alternative assessments. *Research Matters: A Cambridge Assessment publication, 29*, 17–22.

Williamson, J. (2025). How do candidates annotate items in paper-based maths and science exams? *Research Matters: A Cambridge University Press & Assessment publication, 39*, 66–89. https://doi.org/10.17863/CAM.116170