AQA on-screen assessment pilots: insights from focus groups

Research and Innovation

Jeanne Marie Ryan and Corina Balaban

Summary

This report presents research findings from AQA's first series of summative on-screen assessment (OSA) pilots, which took place during April to May 2022 and featured tests in GCSE English Language, GCSE Science (Biology and Chemistry) and GCSE Mathematics.

The objective of the research study was to investigate teachers' perspectives on OSA and the OSA delivery process, and to evaluate students' experiences of interacting with OSA items. Interviews and focus groups were held in respect of each subject and an analysis of the data generated four key themes:

- **Adequate preparation** would be essential in advance of any digital exams being introduced, with students and teachers agreeing that preparation should begin in earnest from Year 7.
- **Concerns about infrastructure** were raised by teachers and by students; they reported that improvements would be needed before high-stakes exams could be delivered digitally.
- **Functionality of OSA** received mixed feedback, with some elements viewed positively (e.g. in English, the facility to highlight text) while others were seen as problematic (e.g. difficulty inputting symbols when responding to maths items).
- **Fairness** emerged as an important concern, with students and teachers keenly aware of the digital divide[1] and the potential for students who have less access to technology and resources to be placed at a disadvantage when it comes to high-stakes digital exams.

The challenges highlighted by the students and teachers closely align with the main barriers that Ofqual (2020) identified in relation to greater adoption of high-stakes OSA.

Context

On-screen assessment (OSA) is by no means a new concept; research studies evaluating the design, development and implementation of OSA have been conducted since the early 2000s (Bennett, 2003; Kingston, 2008; Redecker & Johannessen, 2013). The recent rise in online teaching and learning as a result of school closures during the Covid-19 pandemic has further accelerated the demand for reliable digital assessment. It has also acted as a catalyst for an increase in the use of technology in the classroom (Ofqual, 2020). In England, this has led to renewed policy interest in how technology could be integrated into high-stakes assessment.

The pandemic has exposed gaps in digital literacy, preparedness and infrastructure. According to the Department for Education (DFE, 2021), 46% of state secondary schools did not have an EdTech strategy in place prior to school closures in 2020, and one in five schools considered their

[1] The gap that exists between those who have readily available access to technology and those who do not, and the subsequent impact that this divide can have upon student performance.

existing technology to be inadequate. The pandemic also exposed huge variation in terms of schools' and students' preparedness for digital education (Montacute & Cullinane, 2021). As is to be expected, schools where teachers had been using technology in the classroom prior to the school closures in January 2020 had a much smoother transition to remote education than those who had limited or no technological access.

While digital assessment has already been used across a variety of educational contexts (e.g. vocational education), the focus is now gradually shifting towards high-stakes digital exams for GCSEs and A-levels. In England, the transition to OSA has been prompted by a desire to align assessment practices with the extensive use of digital technologies seen in wider society, to strengthen validity of assessments through more robust assessment design and data collection practices, and to explore new ways to improve teaching and learning (Ofqual, 2020).

In this report, we draw on international examples of OSA practices, predominantly Finland, New Zealand, Israel, Wales and Australia. As these countries are advanced users of OSA, their experiences of digital transition are useful for us to consider as we explore some of the common challenges and identify constructive ways forward.

Motivations for transitioning to OSA have been similar across the world: the desire to modernise the assessment system and to ensure that the evaluation of knowledge and skills is in line with the use of digital technologies in the classroom, in students' lives and in employment (Pearson, 2022). In Finland, for example, one of the key reasons for the shift to OSA has been the need to address the disconnect between low levels of digital literacy in the classroom and high levels of digital literacy in the workplace (Pearson, 2022).

In line with this thinking, the findings of this report suggest that teachers and students are aware of the need to bring assessment into the digital world, acknowledging its 'inevitability' despite the challenges that may be faced in implementing it.

Research design and methods

AQA's first round of large-scale summative OSA pilots were delivered between late April and early May 2022. The subjects included in the pilot were GCSE English Language, GCSE Mathematics, GCSE Biology and GCSE Chemistry, with students sitting a 45-minute test in each subject[2].

For GCSE English Language, GCSE Biology and GCSE Chemistry, two platforms were used to compare the different types of interactive tools offered by each and to gather detailed feedback on students' use of tools and functionalities.

Due to the specific functionality required for on-screen GCSE Mathematics items, only one platform was appropriate for inclusion in the pilot tests. Maths has been identified as the most challenging subject to convert from paper to digital. One of the key challenges is that students are used to doing their workings out on paper. Additionally, questions targeting higher-order maths skills require novel on-screen tools, calculators and functionality, which students are not accustomed to using (Fishbein et al., 2018). We were therefore interested in gathering feedback specifically related to the on-screen input of responses to maths problems.

Our research study involved designing post-test focus groups with students and teachers to:

[2] At certain centres, there were students who took more than one subject, and some took all of the subjects offered.

- gain insight into students' perspectives on taking an on-screen assessment
- gain insight into teachers' perspectives about the experience of delivering on-screen assessments.

These research questions were devised to investigate students' experiences of the functionalities offered by the two platforms used in the pilots and to gather detailed feedback about their experience of OSA at both an item level and more generally. With regard to teachers, investigation focused on their perspectives of delivering OSA, as well as their opinions on what support would be needed if OSA were to be introduced for high-stakes exams and what information reporting tools should provide.

When identifying schools for participation, the aim was to ensure that there was sufficient variation within the sample with regard to average levels of achievement, socio-economic status, school type and geographical location. To achieve this, we used a sampling frame that grouped schools into strata (1–5), with Stratum 1 containing the lowest-achieving schools and Stratum 5 containing the highest-achieving schools. Levels of achievement were based on average performance at Key Stage 2 and GCSE, with data regarding schools' attainment gathered from the 2021 school census[3]. Following stratification, schools were then mapped to their geographical locations across England to ensure geographical distribution.

Due to the ongoing impact of the Covid-19 pandemic, there were difficulties in recruiting the full cohort of sampled schools, particularly those in Stratum 1. As a result, some schools that were not in the original sampling frame but have strong relationships with AQA also came forward to participate. The lack of participation from Stratum 1 schools means that further research is needed to capture more data from schools with lower levels of student attainment. However, it is worth noting that recruiting such schools for educational research is an ongoing difficulty, and this has been exacerbated by the impact of Covid-19 (cf. Armstrong & Finch, 2021).

Twenty-one schools took part in the pilot tests (see Appendix A), of which nine participated in the focus group study.

Stratum	School ID	School type	School location
5	School A	State-funded	South East
3	School B	State-funded	South East
2	School C	State-funded	Midlands
2	School D	State-funded	North West
2	School E	State-funded	South East
4	School F	State-funded	Midlands
2	School G	State-funded	North West
3	School H	State-funded	North East
4	School J	State-funded	Midlands

Table 1 Schools participating in the focus group study

[3] The 2021 school census featured attainment data from 2019 due to the cancellation of exams in 2020 and 2021.

Teacher interviews and student focus groups were conducted virtually via Microsoft Teams in April and May 2022, with 14 teachers and 78 students. The interviews and focus groups lasted approximately 45 minutes; each session focused on one pilot subject, with the exception of GCSE Science, where Biology and Chemistry were discussed together. One-to-one interviews were held with teachers for each subject, with one interview featuring two teachers.

Qualitative data were analysed using a combined method of deductive and inductive thematic analysis. For the deductive analysis, a coding frame was created to identify key themes pertaining to the research questions, which was then used for the process of thematic coding. For example, some of the initial codes included 'reading experience', 'typing experience, 'better on screen', 'better on paper'. For the inductive analysis, researchers identified key themes that arose independently during the course of focus groups and that were not captured in the deductive coding frame. For example, when discussing what worked 'better on screen', new codes emerged such as 'legibility' and 'editing experience'. As a reliability check, researchers first talked through examples of both deductive and inductive codes to ensure agreement, then they separately coded one focus group and compared their coding to confirm majority alignment of the codes.

Findings

The key findings are presented thematically, based on the codes generated by the researchers, and supplemented with participant quotes that were deemed most representative of the findings. Evidence from relevant academic research is also intertwined to complement the student and teacher insights gathered from the data. This helps to situate the findings from our study within the context of academic research evidence and international digital assessment practices.

The four key themes identified by the coding framework were **preparation**, **infrastructure**, **functionality** and **fairness**.

Preparation

Preparation was the most prevalent theme across all the teacher interviews and student focus groups.

Student preparation

Overall, teachers and students were positive about a transition to digital assessment. However, this positivity was dependent upon students having adequate time to prepare for high-stakes OSA, in a similar fashion to how they currently prepare for paper-based exams, with dedicated practice beginning from Year 7 onwards.

> I think if students were taught at a young age to take tests on screen, it would help, like at the start of Year 7 … so students get used to it. (Teacher, School B)

Studies into mode effects – i.e. how the mode of delivery (on-screen or paper-based) impacts upon students' performance – have determined that time spent using a digital assessment platform, usually over a period of a year or more, has a significant influence on mitigating mode effects when students complete high-stakes on-screen exams (Noyes et al., 2004; Backes & Cowan, 2019). A key factor is that when students become more familiar with OSA platforms and item types, this results in a decrease in their working memory load, thereby allowing them to be more effective at responding to the questions themselves (Noyes et al., 2004; Jerrim et al., 2018).

Recent findings by Armstrong and Finch (2021) support this view. In an evaluation of an on-screen GCSE English assessment pilot, students emphasised the importance of 'preparedness'. They highlighted the need for more opportunities to develop their IT skills and familiarise themselves with using keyboards. They also stressed the importance of having more exposure to different types of OSA.

To be able to engage effectively with OSA, all pupils must have a basic level of computing skills, including the ability to type (OECD, 2012). However, typing proficiency alone will not enable students to successfully transition to a digital format. Armstrong and Finch (2021) found that, despite being considered 'digital natives', the current generation of students still needs support and training in using technology for educational purposes, since OSA requires a different type of engagement to that associated with their usual digital activities (e.g. use of social media).

Teacher preparation

Teachers felt that they too would need significant time to prepare for the delivery of OSA. The majority of participants suggested that preparation should begin with a cohort in Year 7 and continue through to GCSE, with adequate support provided throughout for teachers to build their understanding of the functionality of platforms alongside their students.

Teachers called for specific training that would allow them to make best use of on-screen functionalities and give them the confidence to support their students in working with OSA tools. In particular, teachers called for guidance that would clearly delineate the ways in which students would be marked and whether hybrid submissions would be allowed, i.e. whether students could submit any workings done on paper to be marked along with on-screen responses. In the pilots, only some students used paper for their workings out and these were not submitted for marking purposes.

When asked what would be needed to prepare for OSA, many teachers mentioned that it would be easier for them and for their students to navigate one platform, or a few platforms, that had similar functionality and appearance to the platform that would be used for any high-stakes exams. They emphasised that consistency would be key for their students, since the focus of exam preparation tends to be on ensuring that students are familiar with the exam delivery mode.

> Why do we do mock exams? It's so that the students know what it's like to work for that amount of time, on that paper, in that format and so this would have to be the be very similar to that. (Teacher, School G)

This need for consistency throughout the exam preparation process aligns with prior research into why test preparation is essential for students' exam performance. Namely, as mentioned in the above section, test preparation serves to familiarise students with the circumstances of the exam situation so that their working memory will be free to focus on the demands of the questions during the exam (Noyes et al., 2004).

The majority of teachers questioned how students would be graded during the initial transition to OSA. They expressed concern that the first cohorts of students to sit high-stakes digital exams would experience a negative impact on their performance. Teachers were also concerned that the first cohort could be disadvantaged compared to previous cohorts who had taken the exams on paper unless grades were adjusted to reflect the transition to OSA.

All teachers expected high-quality and detailed guidance about new item types as well as extensive support regarding how their pedagogical practice would need to adapt in order for them to be able to best support students encountering novel item types and test formats.

> We would need guidance about if you did allow working on paper, sort of clear guidance for them about that. And maybe it's just saying: You are allowed to work on paper but please be aware it won't be marked. (Teacher, School J)

The teacher quote above describes the need for clear guidance around whether or not paper used for working out may be collected and considered for marking purposes. Teachers and students alike could see the potential benefit of this but emphasised the importance of clarity and also consistency between modes.

The same teacher (from School J) went on to speak about concerns with how OSA would be marked in the initial years after implementation. The majority of teachers raised questions about whether there would be any modelling for mode effects in relation to grading, particularly during the first years of OSA.

> Yeah, just information about how, you know for the first couple of years, how are they going to be graded? What's it going to entail? What sort of coverage are [the OSA exams] going to have for the questions. Just information really! (Teacher, School J)

Concerns about the comparability of student performance on digital versus paper-based outcomes are significant, and the investigation of comparability of performance between delivery modes is an important area of ongoing research. A review by Csapó et al. (2012) outlines that other countries that have implemented high-stakes digital assessment internationally have employed an array of methods to investigate comparability of items between computer-based and paper-based modes. The same review details that the ways in which jurisdictions investigate comparability depends upon the design of individual assessments and socio-cultural factors such as levels of technological access across the student population. Analyses of comparability have thus been developed to suit different assessment contexts on a case-by-case basis, and findings from this international review suggest that a unique process would have to be developed to analyse comparability between modes for GCSEs and A-levels.

A recent report by Building Digital UK (2021) provides a helpful overview of the level of digital preparedness in England across different types of schools, covering teacher and student perspectives. One of the issues that emerged was the stark contrast between private and state schools in terms of how they were able to navigate the digital environment. Prior to spring 2020, many schools in disadvantaged areas had fewer opportunities to use digital approaches and did not have the infrastructure in place to deliver remote learning. At the beginning of the second period of school closures in January 2021, over half of private-school students had technology in place for remote learning compared to only 5% of state-school students (Montacute & Cullinane, 2021). Findings from the same report also suggested that too little attention had been paid to developing teachers' digital literacy skills. Likewise, the DfE (2021) found that 58% of teachers surveyed reported that confidence in skills was a barrier to increased uptake of EdTech.

In Finland, collaborative stakeholder groups developed a large-scale training programme to ensure that schools and teachers were prepared for the rollout of on-screen exams. The programme employed a waterfall approach: instructors trained 50 teachers in how to use the new OSA technology; these teachers in turn cascaded the knowledge to schools across the country via a large number of workshops (AQi, 2021).

Infrastructure

Each of the teachers interviewed raised concerns about the infrastructure available in their school and how this would impact upon their school's ability to deliver OSA to large numbers of students.

The majority of students were also concerned about the availability and/or quality of technological resources in their schools.

> I think some schools won't be ready yet in terms of the number of computers, laptops, facilities in terms of internet access. So, it just depends. Some schools will be ready. For us, it'll take us another two years just to have the right number of classes with computers, laptops. (Teacher, School D)

All teachers and the majority of students reported concerns that some schools would be better prepared than others in terms of access to technology. The points they raised align with Ofqual's (2020) report, which definitively stated that there are extant issues surrounding infrastructure, resources and internet reliability in England, with wide discrepancies among schools depending on factors such as geographical region. To help address this, the government pledged to improve broadband for 3,000 rural primary schools across England in the next three years (Department for Digital, Culture, Media & Sport, 2022).

Other international jurisdictions have attempted to address the specific issue of varying internet reliability in different ways:

- Finland dealt with the possibility of rural power cuts by distributing assessments ahead of time to schools at risk of losing connectivity (Ofqual, 2020).
- In New Zealand, the government committed to ensuring minimum 1GB internet connection across schools, and additional support was also offered to help students improve their typing skills (Ofqual, 2020).
- In Israel, libraries and community centres were used if schools lacked devices with reliable connectivity to enable all students to participate in on-screen exams (Ofqual, 2020).
- In Australia, schools had spacious rooms with additional computers to allow relocation in case of technical problems (Newhouse, 2015).

While the English context may not be identical, these examples demonstrate different approaches to mitigate the digital divide through collaborating with policymakers, schools and trusts, academics and other stakeholders in order to continuously improve digital access.

Functionality

Students' and teachers' comments regarding OSA functionality were mixed, and were often subject-specific in nature. Across all subjects, however, a positive theme emerged: the majority of students said that they were relieved that they would no longer need to worry about their handwriting being illegible and losing crucial marks as a result.

> I would prefer it on the computer […] because some people can't understand each other's handwriting, so the computer just makes it easier to read. (Student, School H)

The following sections detail the main issues that teachers and students raised in relation to each subject.

English Language

In each of the English Language focus groups, students stated that they wanted to be able to highlight and annotate resource materials easily in order for them to refer quickly to key information when formulating their responses. Of the two platforms, Platform A allowed students to highlight

and annotate easily and therefore received the most positive feedback. Students who had used Platform B commented on the lack of this functionality.

Students stated that they wanted to be able to highlight text easily and for the highlighting to remain throughout the assessment, even once they had moved on to other questions, so that they could both complete and review their responses while referencing key information. In line with Armstrong and Finch's (2021) findings, the students stated that they were used to highlighting text when working on paper and wanted to have the same ability on screen in order to feel that they could read and respond in a similar fashion.

> [I would just like the platform] to be easier to use. Maybe making notes and highlighting and stuff [available for us to use on screen], so it's more similar to writing on paper in some aspects. (Student, School C)

Research conducted into the use of highlighting tools (Goodwin et al., 2020) revealed that the impact of highlighting on successful reading comprehension was greater for students reading text on screen than for those reading text on paper. On screen, students tend to highlight smaller amounts of information, suggesting that they use this technique less frequently but more effectively in order to help with their reading comprehension. In contrast, students working on paper tend to highlight larger amounts of texts, thereby rendering highlighted passages less beneficial for reference purposes. Findings from this study, which builds upon previous research in this area, suggest that a highlighting function is essential for on-screen reading comprehension and that it could serve a more important role in successful reading comprehension in digital assessments than paper-based assessments.

Annotation, in the form of making notes in the margin, was often mentioned alongside highlighting as a strategy that students currently employ when reading and responding to questions on paper, and they would like to see this reflected in the on-screen experience.

> I think it would have been useful if there was slightly more space around the actual bit that we were annotating, so we could write outside [the text]. Because when you have a paper copy, you can just write all over it, but we couldn't do that with the online version. (Student, School A)

Research into the use of digital annotation has shown that it can allow for a deeper understanding of text that is presented on screen and it is a key feature that can mitigate for mode effects when students are reading text online for comprehension purposes (Cho & Afflerbach, 2017). In a similar vein to highlighting, annotation allows students to make notes of inferences they have generated between different sections of texts, and is therefore an essential feature to be included in OSA in order for students to apply comprehension strategies successfully.

Mathematics

Out of all of the pilot subjects, maths generated the most concern from students and teachers in terms of the impact of using a digital format. In particular, they were concerned about students' ability to complete their workings out and to respond to questions efficiently using symbols on screen. Students reported that, while they became accustomed to using the on-screen calculator and maths entry tool during the course of the test, they were slower at using these on-screen tools than writing and responding to maths items on paper. Students also reported that they were used to being able to take notes on their test papers and would want to be able to annotate on screen in relation to text, images and graphs, to allow them to respond more efficiently.

> I found it hard to write my workings out, especially on the maths test, because usually, I'm writing and I've got things in different groups. It was hard to do that when I was typing. If I wanted to draw arrows and things that made sense in my head, I couldn't do that on the computer. So it was harder to show what I was trying to work out. (Student, School B)
>
> I didn't know how to find some symbols that I wanted. And then it took extra time. It's just easier to write the symbols than it is to type them. (Student, School C)

For maths in particular, teachers questioned whether a shift to digital assessment would change the nature of the construct being assessed and, consequently, the nature of teaching and learning. Prior research into mode effects in maths performance has shown that students performed better when completing maths items on paper. In particular, items that require a large amount of working out display the most significant mode effects (Hensley et al., 2015). However, other research has also noted that mode effects can vary greatly across maths students of different abilities, with higher-achieving students being less impacted. Some research has found that there is an interaction between the level of demand of a question and students' ability levels, which complicates the way in which mode effects can be calculated. Namely, lower-achieving students who found a question more difficult could appear to experience greater mode effects, meaning that analysis cannot accurately determine whether difference in performance relates to the question or to the mode of delivery. Analysis of mode effects, among lower-achieving students in particular, is therefore complex and is an ongoing area of research (Jerrim et al., 2018).

While teachers believed that on-screen multiple-choice and selected-response items could be used for assessing lower-order skills in maths, there were distinct concerns about how higher-order skills could be assessed in digital format, especially for items where multiple steps of working out, graph work or image manipulation might be required for a student to show how they reached a solution.

> I do have a worry about you know the facility to do problem solving and to do kind of harder mathematics. Like lots of algebra steps on a written screen I think is going to be very hard and so how you build that in to the process I'm not sure. So those higher order thinking skills. You know some of the basic things but you know, we're talking about, a proof or a 'show that' or those sorts of questions. (Teacher, School A)

The concerns expressed by students and teachers are not unique to our study; in large-scale research evaluating mode effects in different test subjects, maths has been identified as the most challenging subject for transitioning from paper to digital (Wang et al., 2007; Fishbein et al., 2018; Jerrim et al., 2018). In response to this challenge, some countries have taken a staggered approach. For example, while Finland digitised its first subjects for high-stakes exams in 2016, mathematics required additional time for development and was not introduced until 2019 (Pearson, 2022).

Science

Overall, teachers and students were positive about the potential for innovation in digital science assessment, while still having some concerns about the use of on-screen scientific calculators. Within the science OSA tests, there was a balance of maths-oriented and text-oriented questions. Students reported that they found typing to be straightforward for many science responses that required short answers (e.g. a sentence or two) or short essays (e.g. a paragraph). For questions that required more mathematical input, students wanted easier ways to input symbols and the ability to annotate items featuring graphs and images.

The majority of students in focus groups suggested that a special character feature could improve their on-screen experience.

> It's sometimes difficult to do special characters, like if you have to square or something. Maybe have like a special characters' function. (Student 6, School C)

> You could have a box where you can do it [enter special characters]. Where you can see all the different options, where you can just click. (Student 7, School C)

Teachers also reported that tools for special characters, scientific notation and symbol entry would be essential for successful use of OSA in science. Similar to the maths feedback, teachers questioned how OSA could change the nature of assessment for science and the types of items that could be used to assess students' knowledge and skills. Science teachers were keen to see more innovative items and were aware of the underlying shift that would be required within organisations such as AQA in moving towards designing digital assessments.

> I think there's a lot of scope, a lot of potential isn't there for … you know showing diagrams really accurately and you know I think there's lots of amazing things you can do on there but I think it's going to take a bit of a mindset switch for you guys in terms of 'We've always written exams in this format and now we've got this online world that we can use' and making that as user accessible as possible. I think this is really key. (Teacher, School G)

Studies that have investigated digital transition have found smaller mode effects for science than for maths. (Way et al., 2016; Jerrim et al., 2018; Soto Rodriguez et al., 2021). Moreover, as evidenced in our findings, research has shown that students and teachers are more positive about the transition to OSA for science than for maths despite the fact that these subjects share many of the same challenges (Fishbein et al., 2018). While further research is required to explore why this is the case, a study by Nikou and Economides (2016) showed that students, and especially lower-achieving students, were more motivated by interactive digital items and showed evidence of improved performance when completing digital items.

Fairness

Students and teachers identified fairness as a key concern when considering a transition to OSA for high-stakes exams. All teachers interviewed recognised that some schools in England would have advantages in terms of digital infrastructure that could enable their students to outperform those from more disadvantaged schools. They acknowledged that while OSA would not necessarily introduce new inequalities, it would exacerbate existing systemic inequalities, further widening the digital divide.

> There's going to be massive economic disparity, which leads to … already a deep educational inequality which exists. And I think that's something that AQA has you know some sort of social responsibility to think about if they are to pitch on screen exams. There's obviously … social good in terms of saved labour costs … but it could also widen the gap when students from private sch… I mean there are some private schools where every single kid has an iPad. (Teacher, School G)

As highlighted in the introduction to this report, the digital divide between different types of schools in England has been flagged as a significant barrier to OSA (Ofqual, 2020). Other jurisdictions, Finland and New Zealand in particular, have also faced issues of fairness and inequality of digital resources during the transition to OSA. Although Finnish society is relatively egalitarian, fairness still emerged as an issue. While the government could not fund the purchase of computers for every student, schools were able to purchase more computers that could be borrowed by students (AQi, 2021). Furthermore, the digital platform Finland chose to use for assessments was

intentionally kept simple so that it did not require a particularly powerful or expensive computer (AQi, 2021).

In comparison, the government in New Zealand ensured that the infrastructure, technology and connectivity allowed for a level playing field across schools (Pearson, 2022). For instance, there was a mix of 'bring your own device' and the use of school-owned devices, with minimum specifications required. In addition, the requirements in place catered to the lowest common denominator (e.g. there was no requirement for a touchscreen to be used; Pearson, 2022).

Conclusions

The four key themes identified from our research – preparation, infrastructure, functionality and fairness – highlight that participants experienced some of the potential barriers with OSA outlined in Ofqual's (2020) report.

Potential barriers to large-scale implementation of OSA		
Theme	**Ofqual (2020) report**	**Student and teacher data**
Preparation	• Essential for students to practise using platforms to become familiar with new software and/or devices • High degree of student input needed during transition • Need to show that OSA matches wider societal changes in technology • Significant engagement activities with key stakeholders required, especially in the rollout of programmes or pilots	• Dedicated practice time needed for students to become confident digital users • Teachers expect guidance about new item types and support for new digital-oriented pedagogical practice • Students voiced that OSA and typing in particular are more relevant to them than paper-based assessment • Students and teachers want to be active participants alongside awarding bodies during the transition to OSA
Infrastructure	• Inadequate ICT provision in schools and colleges • Not enough physical spaces/devices for a whole cohort to sit an exam concurrently • Unreliable internet and local network capabilities • Insufficient qualified staff to support digital transition	• School infrastructure currently lacking for high-stakes OSA, especially in schools with lower overall levels of technological resources • Significant government support and funding required nationally to improve current inequity of technological access

Functionality	• Need to manage the impact of any mode effect during transition to digital • Must redesign or reconsider what should be assessed – adapt assessment of knowledge/skills to formats making best use of OSA	• Awarding organisations required to develop new methods of assessment design for digital purposes • Questions about whether new functionality in OSA could alter the nature of teaching and learning, for better or for worse
Fairness	• Unequal opportunities for students to practise using the relevant software/devices due to differing levels of technological access • Need to develop valid and reliable adjustments for students with disabilities and with differing access requirements	• OSA exacerbating existing social divides (e.g. differentiated access to internet, familiarity with technology, using devices etc.) • Concerns that the first cohort sitting digital exams could be disadvantaged in comparison to previous cohorts

Table 2 Potential barriers to OSA as identified in Ofqual's (2020) report and in our data

However, from our analysis of these key themes, it is essential to note that teachers and students have positive attitudes towards a transition to OSA, including for high-stakes exams, with two main caveats. First, teachers and students stated that there would need to be adequate time to prepare for OSA, ideally from Year 7 onwards, before it could be implemented in high-stakes exams. Second, teachers expressed that government funding would be required to address structural inequalities that exist across schools, in terms of widely varying levels of access to technological resources.

Students' and teachers' perspectives on OSA differed by subject, with students and teachers both reacting most favourably to the tests in English, followed by science and then maths.

- In English, students were most positive about the ability to highlight and annotate texts while responding to questions.
- In maths, students and teachers alike expressed significant concerns about the use of on-screen calculators and tools for scientific and mathematical notation, and teachers questioned whether higher-order skills could be adequately assessed using an on-screen format.
- In science, students and teachers were more positive about the potential for OSA to produce more engaging and interactive item types.

Recommendations and further research

All participants were aware of the current state of the digital divide in England. Teachers in particular drew attention to the need for government intervention and funding to create a level playing field for all schools as a fundamental first step towards OSA implementation. Without such

intervention and improvement to infrastructure at the national level, the issue of fairness raised by teachers and students alike cannot necessarily be addressed (Ofqual, 2020).

The case studies drawn from New Zealand, Finland and Israel provide evidence of successful OSA implementation in high-stakes exams taken at volume in schools and colleges. However, it is important to note that these educational contexts are miles away, both literally and figuratively, from the educational context in England. Ofqual (2020) has acknowledged the obstacles that must be faced in terms of schools' infrastructures, levels of staffing and staff training, and the significant amount of large-scale planning required to implement OSA in England.

Mixed findings from teachers and students about on-screen tools and functionality strongly suggest the need for more research into the subject-specific impacts of item design and platform functionality. Nevertheless, despite concerns about unfamiliar tools, students and teachers were excited about the possibility of innovation leading to more interactive and potentially more accurate assessment of knowledge and skills.

The findings from this first phase of OSA pilots highlight the important role of fairness in teachers' and students' attitudes towards the introduction of high-stakes on-screen exams, suggesting that future avenues for research into perceptions of fairness across stakeholders would be beneficial. Additionally, due to the lack of participation from Stratum 1 schools, further research would be needed to capture more data from schools with lower levels of student attainment.

Ofqual has acknowledged the particular challenges a transition to OSA might pose for students with special educational needs and disabilities. This is an additional area where it will be valuable to undertake further research.

Overall, while obstacles to the implementation of large-scale OSA have been identified, our findings indicate that, with the requisite support and investment, some of these challenges can indeed be addressed.

References

AlphaPlus. (n.d.). *Extended case study: National onscreen adaptive personalised assessments of school children in Wales.* https://www.alphaplus.co.uk/case_studies/Extended-Wales-Case-Study-v2-FINAL.pdf

AlphaPlus. (n.d.). *Accessibility in the Welsh national personalised assessments*. https://www.alphaplus.co.uk/2019/10/31/accessibility-in-the-welsh-national-personalised-assessments/

AQi. (2021). *Digitising a country's examinations: How Finland moved from paper to digital.* https://www.aqi.org.uk/blogs/digitising-a-countrys-examinations/

Armstrong, V., & Finch, K. (2021). *Evaluation of students' and teachers' experience of remote on-screen GCSE English assessment*. AQA.

Backes, B., & Cowan, J. (2019). Is the pen mightier than the keyboard? The effect of online testing on measured student achievement. *Economics of Education Review*, *68*, 89–103.

Bennett, R. E. (2003). *Online assessment and the comparability of score meaning* (Research Memorandum RM-03-05). Educational Testing Service. https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.507.1577&rep=rep1&type=pdf

Building Digital UK. (2022, 24 May). *Connecting gigabit communities and rural schools: The gigabit journey*. https://www.gov.uk/government/publications/connecting-gigabit-communities-and-rural-schools-the-gigabit-journey/connecting-gigabit-communities-and-rural-schools-the-gigabit-journey

Cho, B. Y., & Afflerbach, P. (2017). An evolving perspective of constructively responsive reading comprehension strategies in multilayered digital text environments. In S. E. Israel (Ed.), *Handbook of research on reading comprehension* (pp. 109–134). The Guilford Press.

Csapó, B., Ainley, J., Bennett, R. E., Latour, T., & Law, N. (2012). Technological Issues for Computer-Based Assessment. In P. Griffin, B. McGaw, & E. Care (Eds.), *Assessment and Teaching of 21st Century Skills* (pp. 143–230). Springer, Dordrecht. https://doi.org/10.1007/978-94-007-2324-5_4

Department for Digital, Culture, Media & Sport. (2022, July 1). *Thousands of rural primary schools to get huge broadband upgrade*. https://www.gov.uk/government/news/thousands-of-rural-primary-schools-to-get-huge-broadband-upgrade

Department for Education. (2019). *Realising the potential of technology in education: A strategy for education providers and the technology industry*. https://www.gov.uk/government/publications/realising-the-potential-of-technology-in-education

Fishbein, B., Martin, M. O., Mullis, I. V., & Foy, P. (2018). The TIMSS 2019 item equivalence study: Examining mode effects for computer-based assessment and implications for measuring trends. *Large-scale Assessments in Education*, *6*, 1–23.

Goodwin, A. P., Cho, S. J., Reynolds, D., Brady, K., & Salas, J. (2020). Digital versus paper reading processes and links to comprehension for middle school students. *American Educational Research Journal*, *57*, 1837–1867.

Hensley, K., Rankin, A., & Hosp, J. (2017). Comparing student performance on paper-and computer-based math curriculum-based measures. *Assistive Technology*, *29*, 140–145.

Jerrim, J., Micklewright, J., Heine, J. H., Salzer, C., & McKeown, C. (2018). PISA 2015: How big is the 'mode effect' and what has been done about it? *Oxford Review of Education*, *44*, 476–493.

Kingston, N. M. (2008). Comparability of computer-and paper-administered multiple-choice tests for K–12 populations: A synthesis. *Applied Measurement in Education*, *22*, 22–37.

Montacute, R., & Cullinane, C. (2021). *Learning in lockdown.* The Sutton Trust.

Newhouse, C. P. (2015). Using digital technologies to improve the authenticity of performance assessment for high-stakes purposes. *Technology, Pedagogy and Education*, *24*, 17–33.

Nikou, S. A., & Economides, A. A. (2016). The impact of paper-based, computer-based and mobile-based self-assessment on students' science motivation and achievement. *Computers in Human Behavior*, *55*, 1241–1248.

Noyes, J., Garland, K., & Robbins, L. (2004). Paper-based versus computer-based assessment: Is workload another test mode effect? *British Journal of Educational Technology*, *35*, 111–113.

Ofqual. (2020). *Online and on-screen assessment in high stakes, sessional qualifications.* https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/943382/Barriers_to_online_111220.pdf

Pearson. (n.d.). *Our vision and mission for onscreen assessment*. https://blog.pearsoninternationalschools.com/our-vision-and-mission-for-onscreen-assessment/

Pearson. (2022). *Spotlight on onscreen assessment* (Policy Series No. 5). https://www.pearson.com/uk/news-and-policy/spotlight/onscreen-assessment.html

Redecker, C., & Johannessen, Ø. (2013). Changing assessment—Towards a new assessment paradigm using ICT. *European Journal of Education*, *48*, 79–96.

Soto Rodríguez, E. A., Fernández Vilas, A., & Díaz Redondo, R. P. (2021). Impact of computer-based assessments on the Science's Ranks of secondary students. *Applied Sciences*, *11*, 6169.

Wang, S., Jiao, H., Young, M. J., Brooks, T., & Olson, J. (2007). A meta-analysis of testing mode effects in grade K-12 mathematics tests. *Educational and Psychological Measurement*, *67*, 219–238.

Way, W. D., Davis, L. L., Keng, L., & Strain-Seymour, E. (2015). From standardization to personalization: The comparability of scores based on different testing conditions, modes, and devices. In F. Drasgow (Ed.), *Technology and testing* (pp. 260–284). Routledge.

Appendix A: Table of participating schools

The table below presents the overall number of tests taken between the two platforms. Platform A represents the supplier TAO and Platform B represents the supplier BTL.

System	Centre ID	Centre Name	Number of tests
A	T003	School A	193
A	T004	School B	45
A	T034	School C	67
A	T006	School D	109
A	T035	School E	42
A	T025	School F	99
A	T032	School G	154
A	T026	School H	379
A	T030	School J	49
A	T029	School K	17
A	T031	School L	15
A	T023	School M	26
A	T005	School N	19
A	T001	School O	41
B	CN003	School P	42
B	CN004	School Q	73
B	CN015	School R	93
B	CN016	School S	27
B	CN006	School T	58
B	CN022	School U	147

B	CN017	School V	40