



Story-Based Assessment Pilot

Project report

We're NCFE. We're shaping smarter learning.

We're NCFE: an educational charity and leader in vocational and technical learning. **We combine over 170 years of education experience** with deep insight, working with a network of expert collaborators to shape smarter solutions around the greatest learning needs. In doing this, we're working for a fairer education system for all learners to power inclusivity and choice.

In 1848, we were born from the belief that no learner should be left behind. Today, we're taking up that cause with fresh energy. Our vision and goals will be achieved through:

- Creating education for a fairer world
- Moving towards a smarter education eco-system
- Using our influence to shape real change
- Promoting the idea that potential is personal.



The Really NEET Project was founded to meet the unique needs of NEET (Not in Education, Employment, or Training) young people (aged 16-25) to learn in a supportive and specially tailored environment. They hope to empower positive change in young people by giving them the tools to create stability and security in their lives that will enable them to achieve their future goals and become contributing members of wider society.

Table of contents

List of tables	4
Acknowledgements.....	5
Really NEET and the research team	5
Executive Summary.....	5
Introduction	7
Previous research on games-based assessment	11
The study.....	13
Aims of the study.....	13
Assessment design.....	13
English Reading Assessment Level 1 Examples	17
Maths Calculator Assessment Level 1 Examples.....	18
Sampling the participants.....	19
Data collection and analysis	20
Ethical considerations.....	22
Findings	23
Young people’s experiences of traditional exams.....	23
Young people’s views on whether SBAs would benefit them.....	26
Young people’s views on concentration and distraction	35
Staff views on SBAs.....	39
Limitations	41
Summary of results	41
Conclusion.....	43
References	45
Appendix A: year 1 data.....	48
Appendix B: year 2 data	49

List of tables

Table 1: Participants by setting.....	19
Table 2: SEND groupings by year	20
Table 3: Number of participants who completed each survey by year	21
Table 4: Questions used to determine 1-4 rating	22
Table 5: Overall positivity scale.....	27
Table 6: Participants' positivity rating towards SBAs.....	27
Table 7: Participants' positivity rating by SEND group and Year	28
Table 8: Positivity rating by ASC and SEMH diagnosis	28
Table 9: Positivity rating by gender and year	30
Table 10: Positivity rating by year.....	30
Table 11: Exam preference by year and subject.....	31
Table 12: Confidence level by year and subject.....	32
Table 13: Likelihood of undertaking qualifications with SBAs by year	33
Table 14: Helpfulness of SBAs by year	33
Table 15: Year 1 participants distracted by story/images by survey	36
Table 16: Year 2 participants distracted by story/images by survey	36
Table 17: Year 2 positive and negative distraction levels by subject	37

Acknowledgements

The research team would like to thank the teachers and leaders in each of the educational settings who participated, all of whom showed complete dedication to improving the educational experiences of their students. We would also like to thank the two cohorts of students who gave up so much of their time to take part in the project. We are very grateful for their energy, enthusiasm and unwavering honesty in their feedback.

Really NEET and the research team

This pilot study was designed and conducted by a group of practitioners at Really NEET, a not-for-profit social enterprise that was founded in South Yorkshire. Starting from teaching excluded students out of a room in a warehouse with limited resources, Really NEET now has four sites across England, in Rotherham, Barnsley, Telford and East London. It provides alternative education to young people aged 16-25 from marginalised backgrounds who have struggled in mainstream education. All have an Education Health and Care Plan (EHCP), a cross-agency support plan given to young people with high levels of Special Educational Needs and Disabilities (SEND).

Really NEET's founder, Sophie Maxwell, was motivated to set up The Really NEET Project by her own experience of feeling written-off by school. Struggling with homelessness, domestic violence and undiagnosed learning needs, school repeatedly failed to support her learning despite her significant talents and skills, and she eventually dropped out of school altogether. The frustrations felt by being labelled 'naughty' and unfit for education have been the driving force for everything she has done since; The Really NEET Project aims to ensure vulnerable young people get the support they need to make positive changes despite the challenges they face. With a track record in reducing the proportion of 16–25-year-olds who are Not in Education, Employment or Training (NEET), The Really NEET Project is now a thriving social enterprise, reaching over 120 young people every year.

This report was authored by Laura Quick from UCL Institute of Education with Emma Davies from Really NEET, who led on research design and data collection.

Executive Summary

This study grew out of frustration; the frustration of watching vulnerable young people working hard to gain basic skills and then, when needing to demonstrate this by gaining a qualification, repeatedly failing to do themselves justice. Traditional paper-based exams left them anxious, angry and distressed, and unable to prove their skills and knowledge.

It set about finding out whether a new, innovative form of assessment might go some way to solving this problem. It devised pilot alternative 'mirror' maths and English assessments incorporating Level 1 Functional Skills questions, as an alternative to the conventional maths and English Functional Skills assessments. These were story-based, digital and interactive and piloted with a group of young people across different settings, all of whom had a diagnosis of

Autistic Spectrum Condition (ASC), Social, Emotional and Mental Health (SEMH) or Attention-Deficit Hyperactivity Disorder (ADHD). Young peoples' views and experiences of completing these story-based assessments (SBAs) were then gathered through surveys and analysed alongside their staff's survey responses and researcher observations.

The findings give strong support to the idea that functional skills SBAs could help a substantial number of young people with ASC, SEMH and ADHD. Almost three quarters of our participants felt that SBAs would be of benefit to them, and two thirds preferred them to traditional exams. Many said that this was because they felt less anxious and more relaxed in SBAs. In part because of this reduction in anxiety, many participants reported increased engagement and an ability to focus and concentrate on the tasks. Both the reduction in anxiety and the increase in focus and perseverance also emerged strongly from staff feedback and researcher observations. Importantly, SBAs appear to have the potential to motivate young people with SEND to gain qualifications – responses to the English SBA found 70% thought they would be more likely to engage in a qualification if they were assessed with SBAs while the maths responses reported 52%.

Introduction

Qualifications are necessary for almost all apprenticeships, vocational and educational courses and most employment, and lack of qualifications makes young people's transition into adulthood considerably more difficult. Research from the Children's Commissioner for England (2019) found that 18% - almost one in five - left education without substantive qualifications (that is GCSE grade 9-4 (A*-C) or equivalent) between 2015 and 2019. Moreover, young people with SEND are very much more likely than average to be in this group; the same research found this was the case for an astonishing 45% of 19-year-olds with SEND.

This has serious consequences for their life prospects. Lower educational attainment is associated with lower lifetime consumption and wealth, poorer physical and mental health and negative effects on family relationships (Gladwell, Popli, and Tsuchiya 2022). The ONS Labour Force Survey shows that over the four quarters to Q2 2020, 28% of 16-24 year olds with disabilities were not in education, employment or training (NEET) compared with 8% of those without disabilities (in Powell 2021). The overlapping category of being identified as having SEND also means you are much more likely to become NEET (Gladwell et al. 2022; Powell 2021). Those with significant periods of unemployment following the end of compulsory education have lower long-term participation in the labour market and lower earnings (Gladwell et al. 2022). Too often those with SEND can find themselves in a state of educational and economic exclusion, sometimes in combination with stigma and prejudice associated with their condition; with little money and restricted access to many conventional sources of social interaction and status, this can also be a source of social exclusion and isolation. In addition, any employment may be in lower level, less well-paid jobs than if they had achieved the qualifications of which they are capable.

This under-involvement in the labour market, and associated impacts such as poorer health, involves considerable financial cost to society. Work carried out for the Audit Commission (Coles et al. 2010) looking at the losses to the economy and to individuals and their families of those who are NEET, and comparing this to the cost of youth support, provided striking examples of the financial benefits of delivering appropriate interventions to those with SEND, as well as to care leavers, young offenders and teenage parents. What emerges is a picture not only of individual suffering for those who do not receive the help they need but also of the tremendous waste of the contributions these young people could, over their lives, make to society.

Although information is scarce, the problems caused by the pandemic will also be impacting young people with SEND. Recent research has found a dramatic increase in severe mental health issues in 18-25 year olds since the pandemic (Jackson et al. 2023) and other studies highlight the impacts of educational disruption and loss of employment, family stress, and damage to adolescent mental and physical health (Brooks et al. 2020; O'Shea 2021). These seem likely to affect young people's trajectories as they leave compulsory schooling. Even before the pandemic there were signs of increasing problems; there was a 28% increase in those leaving education without qualifications between 2015 and 2019, reversing a steady drop between 2005 and 2015 (Children's Commissioner for England 2019)

There are several organisations, including Really NEET, providing intensive personalised support to young people with SEND, aimed at helping them move into education, employment or training. This support is in some cases extensive, covering mental and physical health, homelessness, isolation and sexual and domestic abuse. Many of these young people have limited or no educational qualifications (Powell 2021) and this acts as a serious block to accessing jobs, apprenticeships or further education. Helping them gain basic qualifications is, therefore, an important aspect of this support.

Functional Skills assessments were developed to offer an alternative, vocational, pathway to the GCSE route. Functional Skills assessments are currently available in both paper-based and computer-based forms, at five levels - Entry 1 to Level 2, with Level 2 equivalent to GCSE grade 9-4 (A*-C). They can be completed in a wide range of college, alternative, community and work-based settings, with the flexibility to sit assessments all year round. The entry levels are generally assessed internally but Levels 1 and 2 externally by a range of educational companies under Ofqual guidelines. The pass rates vary considerably between companies, arguably reflecting differences between providers and the students they cater for, and are usually significantly lower for Maths than for English (Pearson 2022) – the two subjects considered most important. Successful completion of the higher levels is often required to graduate from apprenticeships and apprentices, often older, have higher pass rates than younger students or those in non-vocational settings (Pearson 2022).

It is not unusual for students to take, and fail, functional skills assessments repeatedly despite their providers often feeling they have done the work and have the knowledge that should enable them to pass. At Really NEET we find that a substantial proportion of the young people on our programmes work hard to master the material needed to pass their functional skills tests successfully, but struggle in the assessment itself. This is partly due to exam stress. There is considerable evidence that test anxiety can hinder academic attainment (Burke, Hryniuk, and Edmondson 2020; Cassady and Johnson 2002; Stöber and Pekrun 2004) and damage wellbeing (Buchanan, Hargreaves, and Quick 2020; Reay and Wiliam 1999), and ‘the higher the stakes, the higher the anxiety’ (Nadeau n.d.). Even high-achieving students suffer anxiety about maintaining their competitive position and the risk of letting themselves down (Keddie 2016), but for our young people this is likely to be a much stronger factor. Almost all have had overwhelmingly negative experiences of school and education; young people at Really NEET invariably talk of disrupted compulsory schooling, poor engagement with school and numerous unsuccessful attempts at accessing post-16 mainstream provision with several different providers. They overwhelmingly see themselves as school failures, with all that this implies for self-esteem and their willingness to re-engage.

Assessments are a particularly stark indication of one’s worth and value in the eyes of others, and for those who perform badly, a source of shame and, often, public humiliation, becoming an aspect of education they may fear the most (Quick 2023). As a result, for many young people with SEND, tests whose format conveys the message that you are being judged in the way that you were previously judged inadequate at school, may create barriers for successful engagement. In addition, a number of those with SEND rely upon reasonable adjustment arrangements (such as readers or quiet rooms) that in many settings don’t materialise or are not sufficient to meet their needs.

Those with SEND may particularly struggle with assessments. A Department for Education study (DfE 2023) found that the average attainment score of pupils with SEND at 16 years old is close to half that of pupils without SEND and, as noted above, 45% of young people with SEND leave school without qualifications (Children’s Commissioner for England 2019). Experiences of failure at school are likely to exacerbate anxiety around testing. ADHD expert Kathleen Nadeau argues that:

Those with ADHD... have a second layer of anxiety to deal with; most come into an exam knowing that their brains have been unreliable partners in the past. Most have experienced the panic of a blank page which produces extra panic for those with ADHD [who] have difficulty organizing their thoughts into a coherent, linear argument. Many others have a long history of frustration when they “go blank” and can’t reliably retrieve information that [they] truly know but can’t retrieve on cue during an exam. (Nadeau n.d.:unpaged)

She points out that students with ADHD have certain cognitive styles that may make test-taking difficult, including difficulties identifying key information, shifting focus between tasks, organising thoughts, cognitive fatigue and having poor time awareness, slow processing speed and working memory challenges.

Similarly, Gardiner (2018) explores how assessments are experienced by students with a diagnosis of autism, arguing that difficulties with social interactions can affect the way students respond to instructions from invigilators as well as their understanding of hidden or subtle meanings within assessment questions, colloquialisms and questions that require a response that involves inference. Moreover, she shows how formal exam settings may be unsuitable for some students, causing sensory sensitivities from, for example, the lighting in the room, perfumes, or the sound of a clock. For those with SEMH, the many issues students frequently have with assessments – anxiety, panic, fear and expectations of failure and low self-worth – may be particularly common, making assessments even more challenging.

There is a long-held belief that summative assessments must be paper-based, or at least a computer-based version of a paper-based assessment, to be valid and reliable. Recently, however, innovations in assessments have begun to challenge this belief, with research beginning to show the benefits of alternative assessments (Molina-Torres et al. 2021). Ofqual’s new corporate plan (2022-25) emphasises the need for innovation and transformation in assessment, particularly in relation to the use of technology, as well as the importance of meeting students’ differing needs in relation to the assessment and qualification process.

Really NEET has always found games-based learning to be beneficial to the young people we work with, both on and off the computer screen. Indeed, it is widely recognised that ‘children with autism spectrum condition (ASC) seem to have an affinity toward digital technologies owing to their linearity and discreteness and often play video games themselves’ (Terlouw et al. 2021:2). Although we had always known how much our learners benefited from games-based learning, it wasn’t until we visited an escape room that we realised there was scope in using games within an assessment context. We watched in amazement as our learners cracked codes and solved puzzles to escape the room within a 60-minute time pressured

environment, not dissimilar to that of a traditional exam. They were able to tackle basic numeracy in the form of multiplication, division and code breaking, showing their working out on the board provided. The young people loved the experience and were relaxed and engaged in the activity, demonstrating skills that intimidated them in a traditional classroom when labelled as 'maths'. Being immersed in an experience that felt safe, relevant, and familiar seemed to be what made the difference. We increasingly, therefore, began to wonder how they would fare if they were able to access assessments in a similar format.

Our intuition that young people like those we worked with would benefit from an alternative form of assessment was supported by further investigation. First, we created paper story-based assessments using storylines such as Zombie Apocalypse. These assessments were taken to our centres and to a mainstream secondary setting so that a small group of learners could trial them and complete a short survey on the experience. Between a half to three-quarters of respondents reacted positively to questions about an increase in concentration and academic results, and a preference for this format over traditional, paper-based exams. We then did something similar with a small group of staff, all of whom were overwhelmingly positive, answering 'yes' when asked whether they thought doing the assessment in this way would improve results and concentration. A deputy headteacher commented, "It's the longest I've ever seen the group concentrate... this is the future of examinations'.

The next stage was to create a physical escape pod, complete with padlocked compartments and a timer to indicate when the assessment had ended. This enabled learners to answer questions to generate a code for the next padlock to be opened. This pod assessment was trialled within the Really NEET project and the response was again hugely positive. However, because the pods cost approximately £500 each, they were not suitable to create for a mass market.



Interactive pod, available to view at Sheffield Hallam University

We then devised the idea for a pilot project examining the feasibility of story-based assessments as an alternative to traditional paper-based exams and pitched this a number of times over the following five years before being awarded a grant from NCFE's Assessment Innovation Fund in 2021. Our aim was to explore whether innovative story-based assessments might have the potential to reduce anxiety and increase both young people's concentration and their likelihood to pursue qualifications. Although our pilot was carried out with young people with SEND, the results might also be of relevance to the wider group of young people who struggle to do themselves justice within the traditional examination system.

Previous research on games-based assessment

Although assessments within mainstream education have generally been slow to adapt to technological advancements, a few initiatives have been introduced in schools and further education institutions, often referred to as ‘technology enhanced assessments’. One review by Timmis et al. (2016) identified a wide variety, from e-portfolios to gaming and even haptics (technology that transmits tactile information using sensations such as touch). The majority of these, however, seemed to introduce technology primarily to improve efficiency of delivery and marking, rather than to aid students, and in general they found that research in this area remains relatively undeveloped. Nevertheless, as well as identifying several practical benefits to these approaches, including flexibility in terms of time and location, they suggest that the use of gaming approaches, including simulations, role-playing and immersive environments, can increase both student achievement and engagement (see also Hickey, Ingram-Goble, and Jameson 2009).

One study by Zainuddin et al. (2020) investigated differences in performance and perceived engagement of science students, comparing those given traditional instruction with paper-based quizzes and those given gamified instruction with gamified e-quizzes. The quizzes were used as formative assessment. Although there were no statistically significant differences in students’ overall learning achievement scores between the paper-based quizzes and gamified e-quizzes, in feedback, students said they felt more emotionally engaged in learning through a game-like system, using words such as ‘fun, enjoyment, interest, enthusiasm, and curiosity’ (p.8). The authors also noted about one of the gaming platforms, ‘avatars, themes, and music are also available... to make learning more enjoyable and interesting’ (p.10) but that students who found it distracting could turn the music off. It may be that the increase in both engagement and control is particularly significant for students with very negative experiences of traditional assessment.

While research into games-based assessments is in its infancy, there are more studies into the benefits of games as part of the learning process. Clarke et al. (2017) describe how the ‘Game-Based Learning’ (GBL) movement started to take on a digital element in 2001, stating that ‘most of the current research that has been conducted concerning the practice of GBL centres on the adoption of various technologies and digital gaming preferences as a means to delivering educational content and exploring techniques of maintaining motivation and engagement’ (p.74). They conducted a small study to test the appeal and educational value of one game – EscapED – within a university setting for a staff training event. The response was hugely positive, with the words; ‘fun’, ‘innovative’ and ‘engaging’ (p. 76) repeated throughout, and all staff feeding back that they could see the value of the game and they would consider using it in their own lesson plans.

A recent systematic review by van Gaalan et al. (2021) explored the evidence for the use of gamification in the education of health professionals. Examining the 44 studies which met their criteria, they concluded that ‘it is possible to improve learning outcomes in health professions education by using gamification, especially when employing game attributes that improve learning behaviours and attitudes towards learning’. These attributes included

assessment attributes, and conflict/ challenge attributes, such as competition and scoring. It is perhaps interesting that studies have found some advantages even with these relatively educationally successful groups; it may be that the benefits would be greater for those with more negative past experiences of traditional exams.

While the primary aim of games-based learning is usually on learning goals, some also involve an element of testing or assessment, often in the form of quizzes used for formative assessment, often self-assessed. For example, in a systematic review, Veldkamp et al. (2020) found that of 39 studies they identified on educational escape rooms, 11 used some form of formative test assessment or evaluation.

Molina-Torres et al. (2021) carried out a comparative study of physiotherapy students' understanding of a learning programme with both a traditional and an escape room evaluation. Each student completed a post-assessment State-Trait Anxiety Inventory (STAI) and Perceived Stress Questionnaire (PSQ), the results of which were compared. Although there was no statistically significant difference in achievement between the two groups, the levels of state-anxiety were higher in the traditional evaluation group, exhibiting statistical significance. Statistically significant differences were also found between the two evaluation systems in terms of the overload factor, the energy factor and the fear-anxiety factor of the perceived stress questionnaire, with the levels of the overload factor and the fear/anxiety factor being higher in the traditional than the computer-based evaluation system.

Taken together, these studies provide evidence that games-based approaches support students' engagement and motivation in learning contexts and can help to improve wellbeing and reduce stress in an assessment context. This comes from both qualitative feedback from students and quantitative self-reported wellbeing scales. This research base, however, remains underdeveloped. Many studies are of a relatively poor quality in terms of sample size and study design, and it is hard to draw clear conclusions about the impact of games-based assessments on students' attainment. Furthermore, in a surprising number of studies participants are university students, or others who have established their ability to succeed in conventional educational terms. It seems likely that repeated experience of school failure, both in learning and in assessments, might make traditional exams particularly problematic and increase the importance of games-based alternatives. For often vulnerable young people, a reduction in anxiety is not only valuable in itself but may also aid their attainment.

The study

Aims of the study

This project grew from a recognition that traditional exams do not suit all students and that there is a need for a more inclusive alternative. To this end we aimed to develop and evaluate one such possibility; that of story-based assessments (SBAs) based on computer games. We chose to pilot these using Functional Skills Level 1 Maths and English as the basis for the assessment, with our sample drawn from young people working at this level who had Special Educational Needs and Disabilities (SEND), focusing on those with a diagnosis of Autistic spectrum Condition (ASC), Social, Emotional and Mental Health (SEMH) and Attention-Deficit Hyperactivity Disorder (ADHD). The project aimed to:

1. Demonstrate that summative assessments can be immersive and interactive while meeting Ofqual criteria and without compromising the standard of assessment.
2. Identify problems some young people experience with traditional exams and see if SBAs have the potential to overcome or reduce these, particularly for those with SEND.
3. Identify the advantages and disadvantages of SBAs, primarily through surveys of young peoples' views and feelings, but also from staff and researcher observations.
4. Determine if the results justify a more extensive piece of research that could use a similar format to construct 'real', externally assessed, assessments that could be taken alongside and compared with the existing format.

Assessment design

The design of the story-based assessments combined the principles of gaming, environmental storytelling (arranging objects so they suggest a story) and escape rooms to create an immersive and engaging experience that aimed to reduce anxiety around the assessment environment and aid concentration, allowing learners to perform at their best academically, enjoy the experience and feel more positive about pursuing future qualifications. The initial design drew on literature in the field alongside our own experience working with young people. The key considerations were:

- **To ensure the game did not distract from the test.** Perhaps most importantly, the storytelling and game element needed to aid completing the tasks rather than distract from them. It was crucial, then, to strike a balance between a meaningful narrative and the assessment questions; the puzzles and functional skills learning goals needed to align (Veldkamp et al. 2020). The narrative also needed to sit within a larger world that the student could see when necessary and that was relevant and consistent with the assessment questions (Nicholson 2016). As Terlouw et al (2021) found when designing an escape room game aimed at facilitating direct communication between

high-functioning children with ASC and their peers for the development of social skills, 'when designing a game with serious purposes, addressing the oxymoron and uniting the fun with the serious aspect is essential but challenging' (Terlouw et al. 2021:3). Key to achieving this is to acknowledge that providing exposition that is not important in solving a puzzle can create red herrings for players. It is thus necessary to tie the challenges into the exposition (Nicholson 2016). The Functional Skills questions needed to be consistently incorporated into the concepts of the setting, world and characters within that world (Nicholson 2016).

- **To use an Ask Why approach.** We adopted the 'Ask Why' approach advocated by Nicholson (2016) in his paper presented to the Meaningful Games Conference in 2016. This involves the designer looking at each element of the player experience and asking 'Why is this here?' (p.4). Each puzzle, task, and item, he argued, should be there for a reason that is consistent with the overall concepts behind the design, such as being part of the setting or world or needed for the character. Doing this is key to environmental storytelling as it makes it a believable experience. For example, a laser maze in Ancient Greece is not plausible and therefore should not be included.
- **To develop a narrative that maintained attention.** Developing an interesting narrative was critical to aiding engagement, focus and reducing exam anxiety. Naul and Lui (2020) argue that 'story is far more than simply an add-on to serious games but an integral piece of the puzzle when creating an immersive, engaging, and motivating learning experience' (p.703). Their review of the literature found that distributed narrative, intrinsically integrated fantasies, empathetic characters and virtual agents, and adaptiveness or responsivity are four characteristics of game narratives that are particularly effective. As Nicholson (2016) puts it, 'players should have a meaningful reason for taking on a task other than 'it's the next thing to do in the room' (p.6). Each challenge needs to have a purpose and be tied into the larger narrative, giving the player a sense that there is meaning to their actions. Alongside story, themes were key to maintaining young people's attention; popular themes were often related to escaping from somewhere such as prison or solving a murder (Jiménez et al. 2020).
- **To expose students to a pregame story about the main character(s).** Learners should be shown a background story before they begin the assessment. Park et al. (2010) found that players who had been shown a pregame story about the main character(s) rated their enjoyment of a game significantly higher than those who had not. This pregame story helps them develop a link to the main characters and therefore feel more invested in finding out how the characters develop within the story. This was important to our design in terms of young people feeling comfortable and that they were not in a traditional high-stakes assessment.
- **To make adaptations to a typical linear structure.** It is common in many games to use a path which follows a linear or sequential pattern - one challenge must be achieved

to access the next. However, this needed adapting to fulfil the needs of the assessment as students needed to be able to skip questions, get them wrong and return to them, whilst the narrative continued meaningfully. Students also could not be given any indication as to the correctness of their answer.

- **To ensure the assessment was as accessible as possible to a diverse group of students.** One aim was to create a platform that, unlike paper-based exams, would seem instantly familiar to students as being a usual part of their cultural world. We knew from observing how our students worked, that the option to include music but also to mute it was important to them and that it was also a common feature of entertainment-based computer games. We also knew that this could at times be distracting, and so opted to offer it as an option in the story sections of the assessment but not the question sections. In addition, as many young people with SEND are allocated a reader in assessments as part of their Summary of Adjustments (SoAs), we wanted there to be not only the option of reading text or it being read aloud, but also to repeat the reading of the question multiple times, particularly valuable for those with poor memory retention. Also, being able to type and point reduces the focus on handwriting, and also allows for deletion and correction, making mistakes less visible. Finally, we wanted to incorporate encouragement to keep going – such as smiles and positive prompts – given that players would not be able to know whether they had completed a task accurately. In traditional exams, this is only available to those young people who have an access arrangement for a prompter (a person who sits with them and refocuses them when necessary) but we felt that this would be beneficial for all.
- **To not attempt to create a one-size-fits all SBA.** We were fully aware that, just like paper-based traditional exams, SBAs were not going to meet every young person's needs and suit everyone. Rather, we were aiming to create an alternative for some of those young people for whom paper-based traditional exams don't work. It was very possible, we knew, that young people who liked paper-based traditional exams or those without a gaming background would not necessarily prefer SBAs.
- **To work towards a more intrinsic motivation for learning.** We appreciated that 'high stakes' summative assessments will always be extrinsically motivated to a high degree, as learners aspire to pass and achieve their qualification, but one of our aims was to encourage greater intrinsic motivation for sitting summative assessments. We wanted the choice of stories to reflect learners' personal interests and therefore offer learners greater self-confidence and enjoyment and the feeling that they wanted to complete the assessment. It was important that the assessments were fun but challenging and that learners were motivated by the personal challenge rather than simply the outcome.

We commissioned a design partner, *Kensa Creative*, to create the final assessments based on these design considerations and their own expertise. All assessments passed through the

same 4 tiers of review implemented for all of NCFE's Functional Skills assessments. The four levels of review were:

1. NCFE initial internal review (Lead Curriculum Officers for English and Maths)
2. Lead assessment Production Officer internal review
3. External Review
4. External Scrutiny

In Year 1 of the project the 4 tiers of review were completed whilst the assessment was in paper format and before they had been sent to the technology partner and online developers. However, in year 2 scrutineers completed the assessment via the online platform created and therefore experienced the assessment in the format experienced by the learners themselves.

After the first year of data collection, new assessments were created and developed based on the students' feedback. Examples of changes made, based on feedback from learners, were:

- The theme for the English assessment changed to a horror theme and the maths to a war theme, based on the top two suggestions offered by the learners.
- More visual prompts were added to encourage learners to keep going. This was due to several learners saying that having someone next to them to encourage them to keep going was beneficial.
- Music and animation were added on request, again, from learners.

Examples from the year 2 Level 1 Functional Skills assessment for English and Maths are shown below:

English Reading Assessment Level 1 Examples



Altrus The writer says that the family crest holds the key to accessing the house from the Entrance Hall. Ah, but you need to find it first!

Altrus What does the word 'crest' mean? Feel free to use a dictionary from the bookshelf, but try not to fall!

Inna

1 mark available

DICTIONARY

crest(noun)

- The top or highest part of something, such as a wave or hill.
- A formal design used by a family, town or organisation as the symbol that represents them.

PREV 1 2 NEXT

Ludwig I heard a click. Quick, push that door open!

Inna Wait, is this the bathroom?

Hanna Water is going everywhere! Inna, what did you do?

Inna I didn't do anything! I just turned the light on!

Ludwig It's another trap. Look at those wires up there, they are still live. If we don't stop the water, we'll be electrocuted. What do you want this time, brother?

PREV NEXT

Altrus Ah, still alive I see. Maybe I underestimated you. But you'll have to do better than that to get out of this one.

Altrus The artwork in this room is a little strange for a bathroom, don't you think? Tell me, what do the images imply?

Ludwig

1 mark available

1 2 3 4 5

PREV NEXT

Maths Calculator Assessment Level 1 Examples



MISSION: 10
Keeping watch
 Available Marks: **3**

To stay alert on watch, you and Cortez both need to consume 25% more calories than normal. Your daily calorie allowance is written on your ID cards.

What are the total number of extra calories needed for both men combined?

SHOW YOUR WORKING **PLAY**

ELIJAH JOHNSON
 FLT OMDR
 HEIGHT 6'3"
 WEIGHT 224LBS
 DAILY CALORIE ALLOWANCE
2800

DANIEL CORTEZ
 B-1 PILOT OFFICER
 HEIGHT 5'6"
 WEIGHT 175LBS
 DAILY CALORIE ALLOWANCE
1850

MISSION: 19
Take aim
 Available Marks: **2**

There are 15 zones in your scope:

- If your shot lands in 1 of the 12 blue zones, you will miss.
- If your shot lands in 1 of the 3 red zones, you will hit.

What is the percentage chance that you will hit the Wendigo?
 Write your answer in your workbook to 1 decimal place.

SHOW YOUR WORKING **PLAY**

ELIJAH
 Where am I?

ICARUS
 Welcome back, Elijah. My name is Icarus, Leader of the rebellion against the Mythias Project. Congratulations! You have completed simulator phase 1.

ELIJAH
 What? I didn't sign up for this.

PLAY

Some questions in the surveys focused on issues related to design and the responses to these aren't explored in this report. This was because the purpose of the project was to see if SBAs might benefit some young people, and so we needed to ensure that the platform was well designed enough to not undermine this, but were not doing a study on design per se. In other words, we needed to make sure that negative feedback was as much as possible due to the format of the assessment rather than problems with the design. The survey responses that focused on design were used to locate problems and improve the usability of the platform for the second set of SBAs, and indeed participants did report increased ease of use in year 2. A higher proportion of participants in year 2 found the platform 'very easy' to use than year 1 – 19% rather than 6% – and a lower proportion found it 'difficult' to use – 13% rather than 19%.

Sampling the participants

Five educational settings were chosen to represent a cross-section of the education provision offered to this age group. This included a mainstream secondary school, a pupil referral unit (PRU), an alternative education setting spread across several sites, a specialist education setting and a 6th form college which changed halfway through the project. The settings chosen were situated in the Midlands and the North of England apart from one in London.

Participants were chosen by the leaders in each of the settings based on two sampling criteria. First, they needed to be taking Level 1 Functional Skills in English and Maths or working at an equivalent level on another qualification. Second, they needed to have one of three diagnoses: Social, Emotional and Mental Health (SEMH), Attention-Deficit Hyperactivity Disorder (ADHD) or Autism Spectrum Condition (ASC). The settings and number of participants in each are shown in Table X:

Setting	Number of participants Year 1 (2021-22)	Number of participants Year 2 (2022-23)	Total
Mainstream Secondary School	3	7	10
6 th Form College	0	13	13
Pupil Referral Unit	6	3	9
Alternative Education Setting	3	27	30
Specialist Education Setting	4	4	8
Total	16	54	70

Table 1: Participants by setting

The sample included 4 females in year 1 and 11 in year 2 (n=15), 12 males in year 1, 41 in year 2 and 2 who didn't specify their gender (n=53). Almost all the participants were white British

apart from five of the year 2 participants: three Asian males, one Asian female and one who identified as Mixed male. Only seventeen participants were 18 or over at the time they took the assessments (4 in year 1 and 13 in year 2), the rest (53) were under 18 and ranged from 15-18 years old.

All participants had a diagnosis of ASC, SEMH or ADHD, sometimes alongside others, and were placed in SEND groupings, shown in Table X:

SEND grouping	Totals		
	Y2	Y1	Y1 and Y2
ASC	20	1	21
SEMH	21	8	29
ASC and SEMH	6	2	8
ASC and ADHD	5	0	5
ADHD, ASC and SEMH	1	2	3
ADHD and SEMH	0	3	3
Unspecified	1	0	1
Total no. of young people	54	16	70

Table 2: SEND groupings by year

Data collection and analysis

Data collection took place between 2021-2023 and involved a set of four surveys completed by two cohorts of participants one year apart, one staff survey and researcher observations.

Before students sat the story-based assessments (SBAs) they took part in initial workshops where they were shown two sample questions in the style of those in the assessments, one from the maths assessment and one from the English assessment and given the opportunity to explore the platform and watch the introductory video. These questions were adapted to ensure they did not match the questions in the actual assessment. This enabled them to become familiar with the platform and the concept of SBAs. Participants were then asked to complete two post-introduction surveys, an initial survey which focused primarily on their impressions of the platform, and one which focused primarily on their experiences of traditional exams and whether they thought SBAs would be helpful or unhelpful to them. The questionnaires were designed and administered on survey planet. There was also verbal discussion that accompanied these online surveys and occasional quotations were recorded by the research lead, but these conversations weren't all formally recorded.

The next stage of data collection involved participants sitting the English and Maths assessment and then completing a post-assessment survey on the same day via a link at the end of the assessments. Almost all learners sat both the English and Maths assessments. If they only did one assessment, this was due to absence on the day. They did not have choice of the assessment sat but sat the assessment delivered on the day that they were present.

Not all participants completed every survey although every participant who sat one of the SBAs completed a post-assessment questionnaire afterwards. The number of participants who completed each of the four surveys in each year is shown below.

	Number of Year 1 participants	Number of Year 2 participants	Total
Initial survey	16 (25%)	47 (75%)	63
Historical experiences of exams survey	16 (24%)	50 (76%)	66
English post-assessment survey	11 (26%)	32 (74%)	43
Maths post-assessment survey	15 (34%)	29 (66%)	44
Total	16 (23%)	54 (77%)	70

Table 3: Number of participants who completed each survey by year

The two post-introduction surveys were mostly multiple choice, with a handful allowing additional free text responses. The post-assessment surveys were identical for maths and English for each year. Again, these were primarily multiple choice with some free text. An extra question was added after the first year to explore distractibility in more detail and to try to establish if the distraction offered was positive or negative to the learner experience.

The data generated by the surveys was exported and analysed using Excel. Because not all young people completed all four surveys, and some young people did not complete all the questions on the surveys they did take, percentages are reported as a proportion of those that answered the question, rather than of the whole cohort. All percentages are rounded to the nearest whole number. Where survey respondents were given the opportunity to write free text answers, these were analysed using thematic coding.

In addition, an overall score was generated for how positive or negative each young person was about SBAs on the basis of the nine questions that covered issues of confidence, engagement, enthusiasm, assessment preferences and attitude to future qualifications. These were all multiple choice apart from one, which allowed for free text responses. The table below includes each of the questions, which survey they were in and whether they were multiple choice or free text. Questions 4-6 were in both the English and Maths assessment and question 2 was only asked of the year 2 participants.

	Survey and question type	Question (multiple choice /free text response)
1	Post-introduction (initial)	Do you think that story-based assessments, like the ones you have seen today, will help you to complete your exams? (multiple choice)
2	Post-introduction (initial)	Do you think that you would enjoy taking an assessment this way? (multiple choice, year 2 only)
3	Post-introduction (historical experiences)	Would having an exam in the form of a story help you and how? (free text)
4	Maths and English post-assessment	Did you feel more confident or less confident going into story-based assessments than a normal exam (multiple choice)
5	Maths and English post-assessment	Did you prefer sitting the online story-based assessment or a traditional assessment? (multiple choice)
6	Maths and English post-assessment	Would you be more or less likely to undertake a qualification in the future, if this were the style of assessment? (multiple choice)

Table 4: Questions used to determine 1-4 rating

Using these nine questions, participants were awarded a number from 1-4 based on how positive or negative their answers were. They were awarded a 1 if they answered positively to all questions, allowing for one ambiguous/neutral/less positive answer. They were awarded a 4 if they answered negatively to all questions, allowing for one ambiguous/more positive answer. A 2 or 3 was awarded for those fairly positive or fairly negative. Where responses varied between the post-introduction and post-assessment surveys, we prioritised the post assessment surveys. All 70 participants in the project were awarded a number from 1-4, even when they had not completed all four surveys.

Ethical considerations

We followed British Sociological Association guidance (2018) on ethical procedures, gaining students' and setting leaders' verbal and written consent, and parents' written consent for students' who were under 18 years old. We provided detailed information (in writing) on the project's aims and process and emphasised to students that participation was voluntary and that they could withdraw at any time without having to give a reason. All names were omitted to ensure anonymity and data stored securely and processed within GDPR guidelines. Learners were selected by setting staff who knew them to ensure they had the capacity to understand what they were consenting to, and our own extensive wealth of knowledge and experience in SEND within various educational settings helped us meet learners' needs well through reasonable adjustments. We also undertook a number of initial visits each setting to build a relationship with learners and leaders prior to the formal research beginning. This initial time investment had a significant positive impact on our working relationships with all

involved and allowed us to develop a strong understanding of learners' individual needs and how best to support these.

Findings

Young people's experiences of traditional exams

As discussed, traditional exams are stressful for many young people but a particular challenge for those who may have a history of school failure with associated low expectations of themselves. This is often also coupled with SEND that can make coping with new or unfamiliar situations difficult or create issues with anxiety and self-confidence more generally, although these may of course be in part the result of a history of school failure. As one participant in the project put it:

I was self-loathing because of my experience with exams and life.

Participants were asked, as part of the 'historical experiences' survey, whether exams worked for them, what support they had been offered, and what improvements could be made. All these questions allowed them to write free text answers, and the responses for years 1 and 2 were analysed together.

Many of the young people expressed a concerning level of distress and anxiety about traditional exams, using phrases such as 'terrified', 'I hate them', 'overwhelmed', 'sick in stomach' and 'I have dark thoughts and feel trapped'. These feelings extended beyond the exam itself and had significant consequences, including not taking the exam:

I'm alright until a few days before and then the pressure of failing causes me to get nervous and stressed.

...at times I can't handle doing them at all because of the pressure and anxiety they cause.

Sometimes I've walked into an exam and straight back out. They cause anxiety because they are big spaces with lots of people, too much paper.

Some explained that their SEND created additional problems, telling us that exams 'tend to trigger my anxiety disorder' or describing their inability to sit still without getting up frequently.

The fear of failure, which was often approached with a feeling of inevitability, loomed large. They described how exams made you feel 'stressed from fear of failing', 'not good at them' and 'like you're gonna fail.' They told us they 'hated failing' and that failure made them feel 'tearful, scared, stressed'. One said:

When I go into an exam I feel shit because I want to do well but don't feel that I can.

While another told us that:

I don't enjoy sitting them and have never done very well. I feel pressured and just want it to end.

Some seemed resigned to this situation:

My mind goes blank in any exam. I'm used to it, it always happens and I'm used to failing.

This feeling of being unable to perform under pressure was reported by twenty-five participants. They said that 'exams haven't always shown an accurate representation of my knowledge' and that they could not think or work as usual in an exam, explaining that 'once I get an idea I lose it', 'I just see lots of words' and 'I'd re-read [the questions] 20 times'. Comments included:

... my mind has gone. I see it but just can't work it out. My mind gets wrapped up in any thought it can think of.

In the exam I find my mind goes blank and I can't remember.

... my brain would go into thousands of conclusions and be overwhelmed.

Concentration and focus could be a particular problem:

I get off task/stare into space. I need help to stay on task because I can't sit still and I'm not allowed to move in exams.

Time was a particular pressure. For some, they worked slowly and needed longer:

... ideas come into my head but I'm slow to write them down.

I had extra time but it still wasn't enough time.

For others, it was the pressure of knowing time was limited that was their main worry:

I am wary of time and feel like I am running out of time.

While a few suggested longer to do the exams, more said how good it would be not to have time limits as it would reduce pressure and therefore aid concentration.

The fact that 'everyone takes the exam together in large spaces' was a particular source of anxiety, and requests for quiet, privacy and smaller rooms were mentioned by twelve participants. Large halls with many people, and the feeling of many eyes upon you, could be intimidating and made one participant:

... distraught because back when I was in school they'd put so many kids in one big hall. You're already worried about the exam and then you worry what your peers are thinking on top.

Many of the suggestions made about improving exams were around overcoming this. There were requests for a 'smaller room to reduce anxiety', 'a small dark room', 'just a quiet space', 'a calmer space' and a 'quieter environment'.

An important aspect of this fear of large halls was that of being watched and judged by others; one complained about feeling 'on the spot and focus on me' and another remarked:

You worry about the people next to you. I worry about people seeing me struggle.

Feeling that others were supportive was important. Suggestions included 'someone saying 'well done'', and feedback that focuses on what you did well. For one, the solution was 'a quiet room with just a few friends taking the exam at the same time'.

However, there were no 'one size fits all' solutions; one participant wanted someone to sit beside them for support and another to be alone in an individual soundproof room and just checked on at intervals. Similarly, while many longed for quiet and calm, others found the silence of the exam hall a problem:

I can't concentrate because they're silent.

I found [exams] awful... because they were quiet. They used to get me stressed.

A bunch of people sitting in silence is not a good environment as I can hear my blood running.

Thirteen young people requested music:

I feel like a small cubicle with just a laptop and maybe a CD player where pre-selected music is loaded on to stop you cheating.

Music choice (MP3 player) and lighting choice.

Three mentioned visuals, one saying that 'I think pictures might help me stay awake' and one who wanted 'videos to distract'.

Preferences between paper-based and computer-based exams were similarly varied. Eleven expressed a preference for working on computers or laptops, often because writing made their hand hurt.

My hand and wrist get tired.

I can cope with maths on paper. I struggle with handwriting so longer answers are hard.

I get tired and can't make out the words on paper based exams whereas computers are HD [high definition].

More generally, paper was seen as old-fashioned:

People my age walk into an exam and just see paper and can't be arsed.

However, seven out of the sixty-six young people who completed the survey were positive about traditional exams, several explicitly mentioning their preference for paper-based exams over computers:

I don't mind paper. I actually did a maths exam on the computer which was more confusing.

I like to write instead of type.

I like writing the information down on paper.

A further five were unsure how they felt about them or had no preference:

I'm not bothered about exams... I do them and then just move on. It's stuff you can't control and so you just have to do it.

Participants' comments about their experiences of traditional exams point to several conclusions. First, many of the young people associated exams with concerningly high levels of anxiety and failure, levels which one might expect to reduce their attainment in exams as well as their wellbeing. Second, when asked what might improve their exam experience, young people suggested different, often incompatible, recommendations. There were no simple, one-size-fits-all, solutions. What was clear, however, was that many young people desired significant changes and most seemed clear of changes they felt would benefit them. Third, 71% of the participants said they had previously been given a Summary of Adjustments (SoA) for exams. Only 16-24% said they had not got a SoA or weren't sure. Although no doubt the adjustments they were given could have been improved, they were often extensive, ranging from extra time to a reader, rest breaks and a separate room provided. This suggests that the problems many young people face in exams cannot simply be remedied by extending their SoAs. These findings supported the need for developing alternative methods of assessment as well as providing information about particular problems these alternatives needed to address.

Young people's views on whether SBAs would benefit them

Given the experiences of many young people with SEND, the impact of the SBAs on confidence, engagement, enthusiasm, assessment preferences and attitude to future qualifications were particularly important to the project. Nine of the survey questions, including one that allowed for free text responses, covered these areas, three from the post-

introduction surveys and three from each of the post-assessment surveys. To capture the participants' overall level of positivity or negativity about the SBAs we constructed a 1-4 scale:

1	Very positive
2	Fairly positive
3	Fairly negative
4	Very negative

Table 5: Overall positivity scale

Each participant was then awarded a number from 1-4 based on their answers to these nine questions. How this was done and the survey questions asked are explained above. At times we also found it helpful to combine the positive responses (1+2) and negative responses (3+4).

The overall results were striking. Of the 70 participants, almost three-quarters were positive and over half felt SBAs would be very positive for them. Just over a quarter were negative. The number of participants rated 1-4 are shown below:

	1 (very positive)	2 (fairly positive)	1+2 positive	3 (fairly negative)	4 (very negative)	3+4 negative
No. of participants	38 (54%)	13 (19%)	51 (73%)	13 (19%)	6 (9%)	19 (27%)

Table 6: Participants' positivity rating towards SBAs

We then broke this down by SEND grouping, ASC/SEMH/ADHD diagnosis, age and gender, as well as separating out the results for year 1 (n=16) and year 2 (n=54). Table 7 shows participants' 1-4 positivity ratings towards SBAs by SEND grouping:

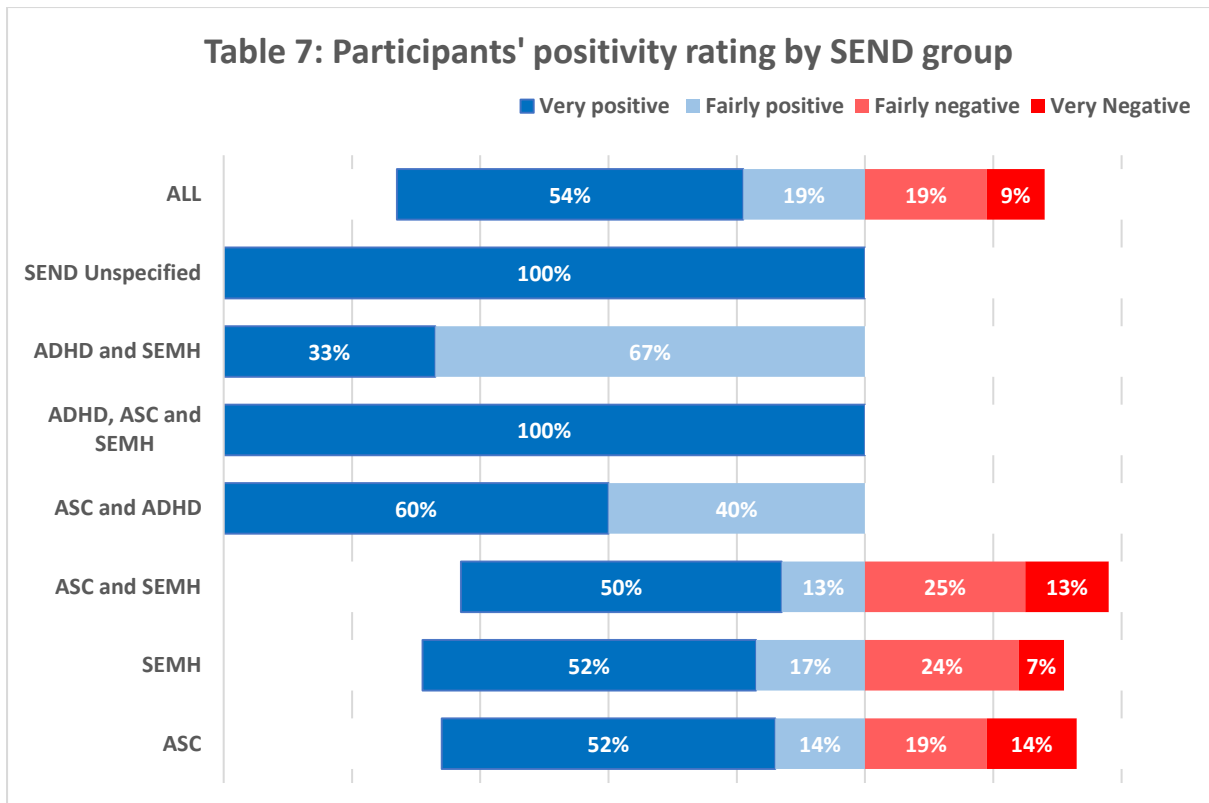


Table 7: Participants' positivity rating by SEND group

We then grouped again into three groups: those with a diagnosis of ASC, those with a diagnosis of SEMH (a number of whom had both) and those with a diagnosis of ADHD (all of whom had either ASC or SEMH). The table below shows the number of participants given a 1-4 rating by year and diagnosis:

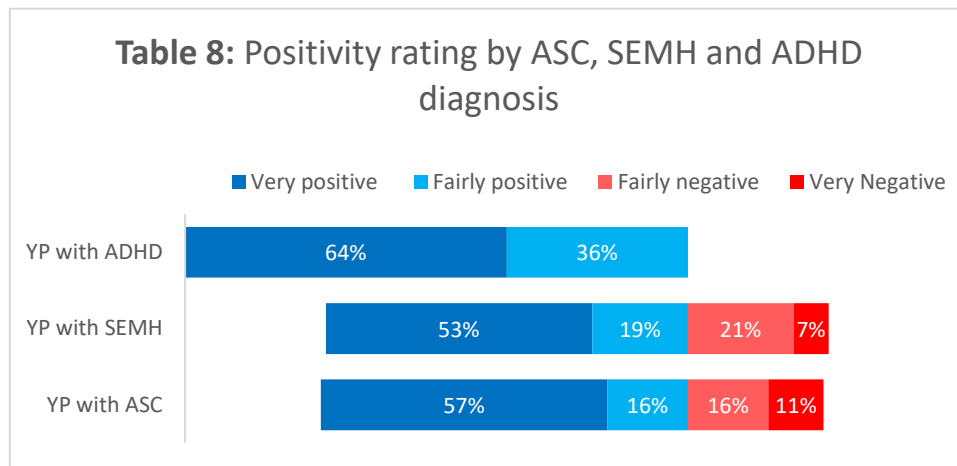


Table 8: Positivity rating by ASC, SEMH and ADHD diagnosis (both year groups)

We had thought there might be differences between the ASC group and the SEMH group but the results were strikingly similar. 73% of those with a diagnosis of ASC felt positive and 27% negative towards SBAs, while for SEMH the figures were 72% positive and 28% negative.

Only eleven had ADHD included in their diagnosis, but it is suggestive that all reported feeling positive. This could indicate that SBAs overcome some of the difficulties those with ADHD face, such as organisation of thoughts, focus and retrieval of information (Nadeau n.d.:unpaged), as discussed in the introduction. Our survey data couldn't shed light why this was found, which should be a topic of further investigation.

We also divided participants into two groups: those under, and those over, 18 years old at the time of taking the assessment. Again, there was little difference between the groups, roughly two-thirds reporting feeling positive and one-third negative in both age groups. We did not analyse by race or ethnicity as only five participants were not white British. Finally, we divided our sample by gender, although the heavy preponderance of males (n=53) allowed us only a suggestion. 86% of females were positive about the SBAs (67% very positive and 20% fairly positive) compared to 70% of males (53% who were very positive and 17% fairly positive). Table 9 shows participants given a 1-4 rating by gender and year:

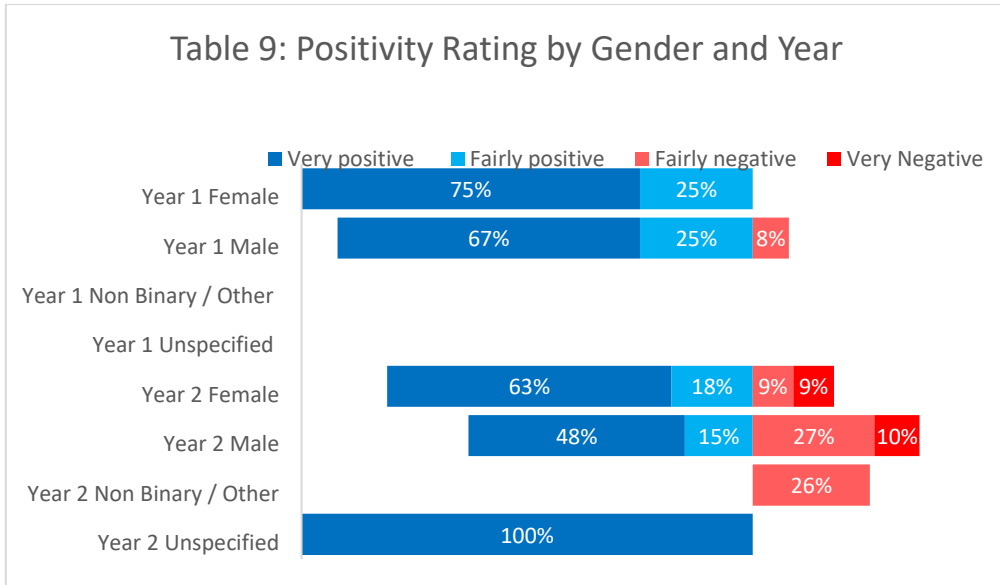


Table 9: Positivity rating by gender and year

The only explanation to emerge from the data for girls’ higher level of overall positivity was the story element of the assessment; a higher percentage of positive girls than positive boys cited the story as the reason they felt positive about the SBAs, a point we explore in more detail in discussion of those questions that asked specifically about story, below.

Finally, there was a marked difference in positivity rates between the year 1 cohort (n=16) and the year 2 cohort (n=54). In year 1, every participant but one was positive about the SBAs, in contrast to year 2, when two thirds were positive and a third negative. The percentages of participants who were awarded 1-4 in each year are shown in Table X:

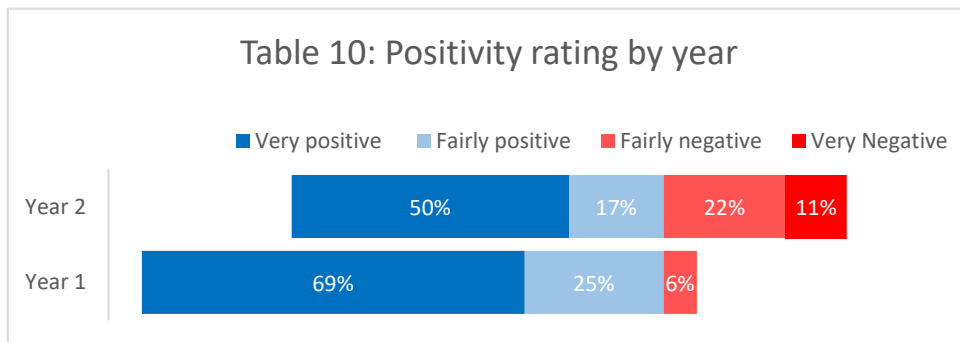


Table 10: Positivity rating by year

This could be for several reasons, including the small sample size in year 1. However, two further sets of survey responses suggest that it may have been due to differences between the popularity of the English assessment story and images in year 1 and 2. When participants were asked if they found the storyline for the English assessment interesting, 100% of Year 1 participants replied ‘yes’ or ‘yes, very’, compared with 59% in Year 2. Moreover, 82% felt

positive about the images in the English assessment in year 1 compared with 63% in year 2. These differences between the years were not, however, found for the maths assessment.

Researcher observations in year 2 offered some explanation of why the year 2 English assessment may have received a less positive response than it did in year 1. First, students found the introduction to the English assessment too long. Verbal feedback from students suggested that some simply became frustrated with the amount of storyline that preceded the first question and some just clicked through the assessment to get to the end due to frustration. Second, there was also a suggestion that for some there was not a clear enough demarcation to indicate that a question was being asked. These findings will be considered if further research is carried out.

It is likely, in short, that the popularity of the year 1 English assessment is responsible for this difference in responses between years. Interestingly, researcher observations included a comment that the year 1 English assessment ‘storyline (Spellbound) [is] too young and feedback suggests this is linked to the colour scheme and visuals’ and two responses in the staff survey from year 1 supported this concern. The year 2 English storyline, based on a horror theme, aimed to overcome this but in fact proved less popular. It may be that the year 1 participants were less willing to make positive responses about a storyline that could be viewed as ‘younger’ – more feelings-based and with cartoony graphics – in discussion with their staff and the researcher than they were in a private survey.

Looking at the nine questions we used to construct our positivity scale in more detail allowed us to develop a fuller picture of how participants felt. When asked whether they preferred SBAs to traditional exams, two thirds agreed, 65% for English and 68% for maths. About a third said they preferred traditional exams, reinforcing the finding that there is no one-size-fits all solution to the problem of high-stakes assessments. Table 5 shows the number of participants who said they preferred each assessment:

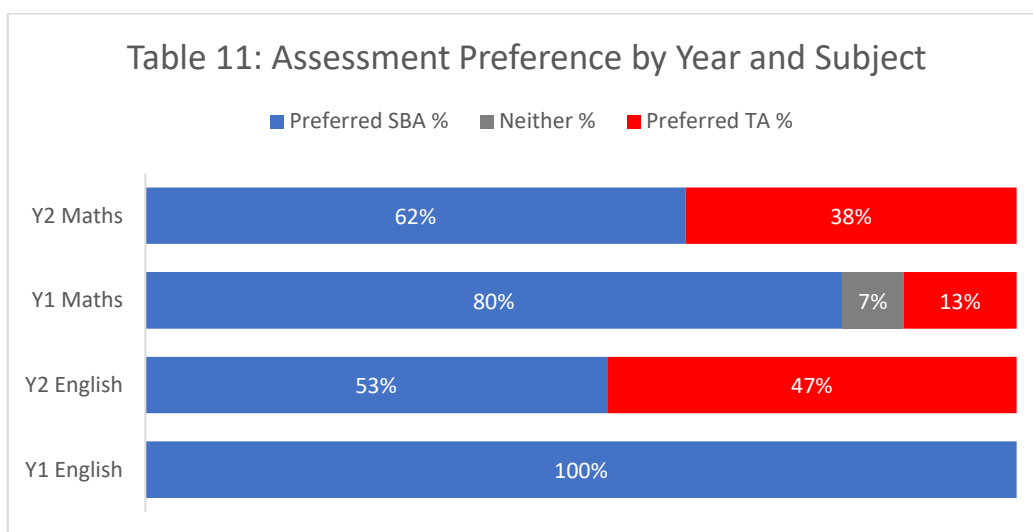


Table 10: Assessment preference by year and subject

When these results were divided by year, the findings mirrored the difference discussed between the year 1 and 2 participants. 100% of those in Year 1 preferred the story-based English assessment compared with 53% in Year 2, perhaps also due to the popularity of the Spellbound story and images. The difference between years was less striking for maths, however. 80% of Year 1 participants preferred the story-based maths assessment to a traditional exam, compared with 62% in Year 2.

In their post-assessment surveys participants tended to report that they felt more confident going into the SBA than a traditional exam. Table 12 shows participants' reported confidence levels by year and subject:

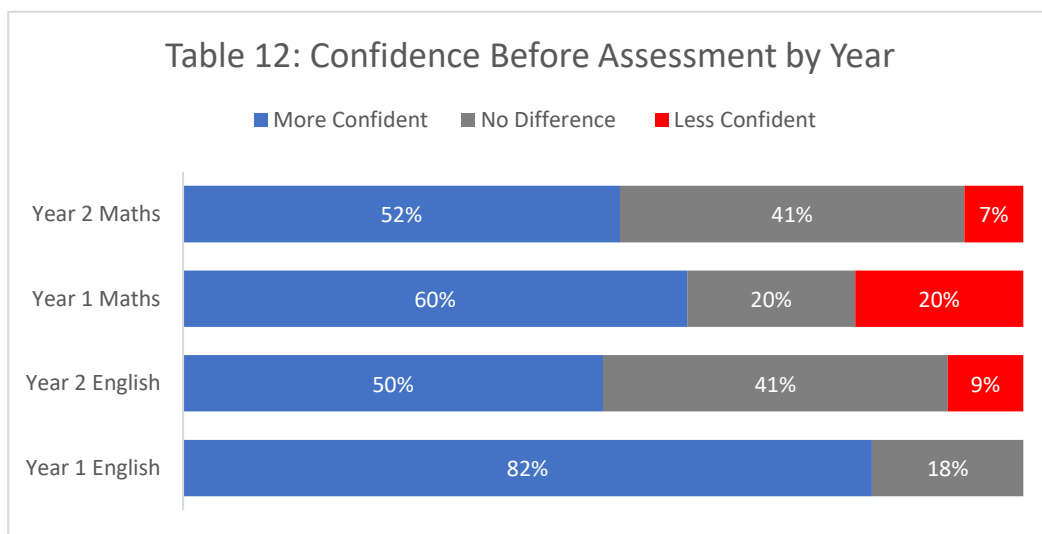


Table 11: Confidence level by year and subject

Although year 1 participants, particularly in the English assessment, were more positive than year 2 participants, the overall figures were roughly similar for maths and English, with a little over half feeling more confident with the SBAs, a third saying there was no difference in their level of confidence and a handful saying they felt less confident going into the SBA. This is potentially an important finding because so many young people with SEND can struggle with confidence due to repeated experience of failure; developing assessments which are less intimidating could increase their motivation and engagement significantly. As noted in discussion of limitations below, many factors may have fed into these findings, ranging from their knowledge that the assessment 'did not count' as it was not their 'real' high-stakes assessment, to anxiety about attempting something new.

Those with a history of exam and wider school failure are often hard to motivate to attempt qualifications, even if they are important to enable them to achieve their goals. We were therefore interested to know if participants felt SBAs would influence their attitude to pursuing further qualifications. In the post-assessment surveys, participants were asked whether they would be more or less likely to undertake a qualification in the future, knowing

that the style of assessment would be story-based. Table 13 shows their responses in both the English and Maths surveys:

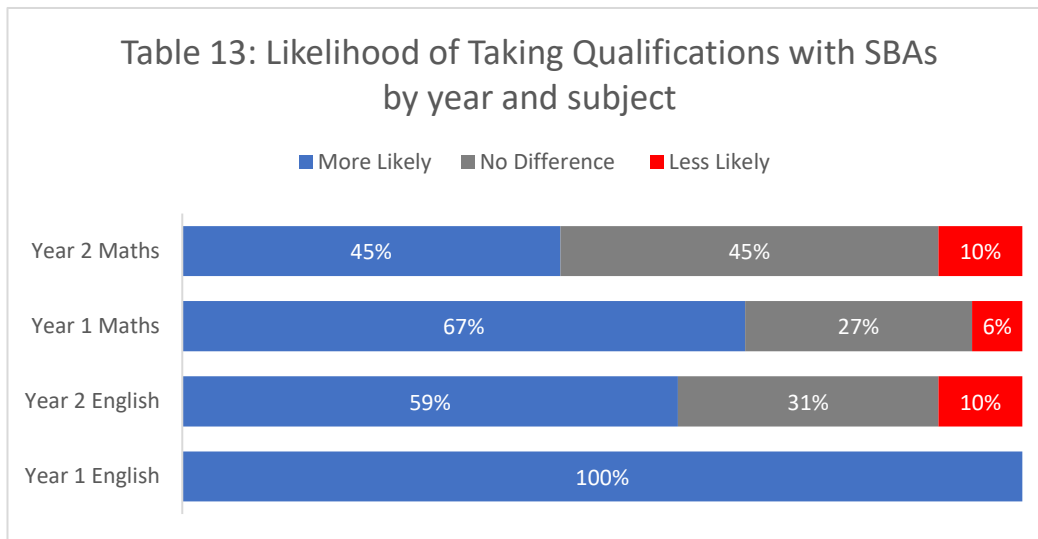


Table 12: Likelihood of undertaking qualifications with SBAs by year

Again, the results were largely positive. In the English post-assessment survey 70% of participants said they would be more likely to take a qualification, with year 1 participants particularly positive. The figures were lower overall for maths, 52% in total, and in year 2 the same number said it made no difference as felt positive. Very few said they would be *less* likely to take a qualification in SBA format.

In the pre-assessment surveys, taken after their introduction to the platform, storyline and characters, participants were asked whether they felt in general terms that SBAs would help them complete exams and why. Roughly two thirds felt they would, with around a third saying that they would not or did not know. Four did not complete this question. Perhaps surprisingly given year 1 participants' responses to several other questions, rather more year 2 participants said that SBAs were helpful. Table 14 shows whether participants thought SBAs would be helpful to them by year:

	Helpful (a little + a great deal)	Not helpful (not much + not at all)	Don't know	Total
Y1	10 (63%)	1 (6%)	5 (31%)	16
Y2	33 (72%)	10 (22%)	3 (13%)	46
Total	43 (69%)	11 (18%)	8 (13%)	62

Table 13: Helpfulness of SBAs by year

To explore this further, participants were asked to give free text responses explaining their answers. These showed up four main themes: a preference for computer or games-based

assessments over paper-based exams; a liking of the story-element of the SBAs; a reduction in anxiety and pressure; and an increase in concentration and focus.

First, the computer-based nature of the assessments proved popular. Three participants commented on a preference for typing over writing, one saying typing was quicker. This supports the many complaints that writing made hands hurt in the 'historical experiences of traditional exams' responses. Two participants said they also enjoyed 'clicking things'. While traditional exams administered on computer could offer these, three also mentioned the benefits of the music and visuals on the SBAs:

... the music helps me to focus better.

... it engages people in a more visually enriching way... It makes it more fun and exciting.

Two also appreciated the silence that headphones offered, one pleased to 'stop the muttering in the corner of the room'. Others mentioned the similarity to computer games, and one appreciated the sequential nature of the questions, remarking:

I can't see every question in one go and so there is not too much information.

A second theme was that the story element made the assessment more enjoyable:

More useful having a story it kept me entertained

... the drama makes it interesting

... a storyline would make it more interesting.

Having everything story based makes it more entertaining/fun

Stories also aided engagement. Fifteen mentioned this. Comments included:

I think it would be interesting so I would want to get to the end

... it feels like you're in the story.

... it will pull me in.

... makes you want to answer the question/influence the outcome.

Across both years, story was the most common reason given for the SBAs being helpful, with only one saying they didn't like stories. Girls were particularly likely to cite story as the reason they felt the SBAs would benefit them, and several of the most positive mentions of story came from girls:

I loved it and wanted to know what happened in the end.

It uses dialogue which helps and I like the character Erkle he's cute.

... it helps as everything links together and makes sense. Having everything story based makes it more entertaining/fun & makes you want to answer the question/influence the outcome.

The attractiveness of the story element to girls was the only reason the data suggested for their slightly higher positively rating towards the SBAs, as noted. It may be that the themes or stories chosen for the pilot assessments were more suited to girls, or that girls benefit most from assessments being story-based.

The third theme to emerge from the free text responses was the reduction in pressure young people felt the SBAs offered. Three were explicit about SBAs feeling less stressful than traditional exams, which they felt was helpful, making comments such as:

... you don't have the pressure like with the paper based exam.

... it helps me de-stress.

This reduction in anxiety was a major theme in staff responses, discussed below, and also researcher observations, which included comments on one participant's reduced verbal ticks and two participants reduced bodily movements:

Two learners rocking backwards and forwards with anxiety about interactions with new people observed physically reduce the amount of body movement when they were interacting with our exam platform.

The final theme to emerge from participants' answers to this question was the impact SBAs had on their concentration. Because this is often cited as a key concern in relation to new and more innovative games- or story-based assessments, these responses were examined alongside others that shed light on participants' experiences of whether SBAs helped or hindered their concentration, and thus likelihood to achieve well in the assessments.

Young people's views on concentration and distraction

One of the concerns raised about computer-based or games-based learning and assessments is that lively, colourful and potentially absorbing characters, stories and game-elements can distract from the educational element. Some fear that learners, in these contexts, will focus their energy and attention away from the tasks and onto these extra features. We therefore asked participants about this directly and included one question on whether they found the story distracting and one on whether they found the images distracting in both the post-assessment surveys and one of the post-introduction surveys. The year 1 survey results showed that although two thirds said they did not find the story or images distracting, a small number did, ranging from 13%-31% across surveys, with the exception of the story in the

English assessment, which no young people said was distracting. The table below shows the year 1 participants' responses by survey:

Distracting?	Initial		English		Maths	
	story	images	story	images	story	images
Yes	3 (19%)	5 (31%)	0 (0%)	2 (18%)	3 (20%)	2 (13%)
No	10 (63%)	11 (69%)	11 (100%)	8 (73%)	10 (67%)	10 (67%)
Neither/other	3 (19%)	0 (0%)	0	1 (9%)	2 (13%)	3 (20%)
Total	16	16	11	11	15	15

Table 14: Year 1 participants distracted by story/images by survey

Our initial assumption was that, for those who reported being distracted, this would likely impact their ability to focus on the task, potentially reducing their engagement with the content of the assessment and therefore their attainment. However, when these responses were analysed alongside their free-text responses to other survey questions, we realised that this was not the only interpretation. For example, one participant responded that 'I'd forget I was doing an exam and concentrate more'. Distraction seemed to have been interpreted by some in our year 1 cohort as a 'good thing' because it distracted from the fear and anxiety that got in the way of them engaging with the assessment tasks. This same participant had described their feelings about traditional exams as 'frustrating, confusing, stressful, hard to remember things and feels like they're trying to trick you'. An alternative assessment, then, able to distract from such feelings, might for some have a positive impact on their ability to focus on the task in hand.

In year 2 we repeated the original questions, finding slightly higher levels of distraction but still a substantial majority claiming not to be distracted:

Distracting?	Initial		English		Maths	
	story	images	story	images	story	images
Yes	10 (21%)	13 (28%)	14 (44%)	8 (25%)	8 (28%)	4 (14%)
No	30 (64%)	28 (60%)	16 (50%)	18 (56%)	19 (66%)	22 (76%)
Neither/other	7 (15%)	6 (13%)	2 (6%)	6 (19%)	2 (7%)	3 (10%)
Total	47	47	32	32	29	29

Table 15: Year 2 participants distracted by story/images by survey

However, we added an additional question to the year 2 post-assessment surveys asking whether participants felt the distraction was positive or negative. In response, half claimed they had been positively distracted, twice as many as said they had been negatively distracted. Table 17 shows their responses by subject:

	English	Maths	Total
Positive distraction	16 (50%)	16 (55%)	32 (52%)
Negative distraction	8 (25%)	8 (28%)	16 (26%)
Neither/neutral/both	8 (25%)	5 (17%)	13 (21%)
Total	32	29	61

Table 16: Year 2 positive and negative distraction levels by subject

Whereas only roughly a quarter said they found the stories or images distracting, when asked the additional question - whether they found the assessments positively or negatively distracting - around half identified as positively distracted and a quarter as negatively distracted, meaning three-quarters claimed some sort of distraction.

We think it is likely that this difference was because many participants initially assumed being distracted was negative, but when asked explicitly whether they were positively or negatively distracted, shifted this perception. This was not true of everyone, however; for example, fourteen young people said the year 2 English storyline was distracting whereas only eight said they found being distracted in this assessment negative, suggesting that six of the initial responses were referring to positive distraction. The concept of 'distraction', then, should be treated with caution or perhaps avoided and replaced with discussions of focus or concentration.

Claiming you are positively distracted in an assessment does not of course mean that this distraction will aid your engagement and attainment in the assessment; participants could have felt that being distracted was positive because it means you *don't* have to focus on the assessment but can forget about it. However, this did not seem to be the case from the responses they made to other survey questions. Like the year 1 participant above who said 'I'd forget I was doing an exam and concentrate more', others also connected it with increased focus:

I think it would help because it stays on one subject and keeps it on track. It feels more organised / better to remember. Takes mind off anxiety.

Some felt that because SBAs were more interesting, they would aid motivation and concentration:

I think it would be interesting so I would want to get to the end

... they would help me to concentrate, especially if I'm interested

A few felt it would aid concentration by lessening the boredom they felt in traditional exams:

... they would be more intriguing and stop me from falling asleep half-way through my exam.

... it would help with concentration as it's less plain and boring.

Responses from staff – discussed in full in the following section – supported the participants' views, as did the researcher observations. These commented that the desire to find out what happens in the story 'appears to distract from the time element and keep learners focused', the researcher being struck at participants' 'ability... to stay seated and focus without need for rest'. One comment noted that 'learners entering the room anxious and distracted and once they are into the storyline, they focus more'.

In contrast, responses from four young people included explicit comments about finding the SBAs *negatively* distracting. These included 'too confusing', 'it could complex the question' and 'a story splits your attention'. Although one felt that although 'a story would overcomplicate things' they would still prefer it to a traditional exam, another had concerns that some stories might be more distracting than others:

... a horror story could help as long as it's not too immersive. It could be more of a distraction.

In summary, participants' responses to questions about their experiences of traditional exams showed that they found concentration and focus a key issue in traditional exams, something often linked with their anxiety and a feeling of pressure. This was aided or hindered by various factors such as noise, silence and the environment. What we learnt from their responses to questions about distractibility was that if alternative assessments can be developed that reduce these by distracting away from their anxiety, for example, with an interesting storyline or the addition of music, they may well aid the concentration of some young people. When an 'undistracted' young person in a traditional exam reports that they are 'terrified', 'overwhelmed', 'sick in stomach' and has 'dark thoughts and feel trapped', it may be that some distraction is what is needed.

Clearly, the story, images, music and characters in SBAs have the potential to have a negative impact by distracting young people from the assessment questions, or to have a positive one, in that by making the assessment more engaging and enjoyable they reduce anxiety and aid focus and motivation. Indeed, we found evidence of both in our small pilot study. Comparison between students' focus on the assessment tasks in traditional exams and SBAs should be an important part of any future research.

However, our participants' responses strongly support the idea that replacing traditional exams with SBAs could lead to an overall reduction in anxiety and an increase in concentration and focus for a substantial number of young people. These findings were supported by the staff survey responses, which we now move on to discuss.

Staff views on SBAs

Seventeen staff across all settings completed a 7-question survey. The majority of the staff feedback was gained in Year 1 but two staff from the new 6th form college setting offered their feedback during year 2 of the project. All staff said they were positive about the potential of SBAs for the young people they worked with, many extremely positive:

Think this is an amazing project, much more inclusive than old pen and paper version.

I fully support alternative methods of exams and assessments like these as it provides options for learners. I think the education system has needed something like this for a while now, so it is exciting.

I am an ex-student [from this alternative provision] and now I am a member of staff. I like the assessments and they would have worked for me.

The whole thing is user-friendly, accessible for all and really engaging all together.

One explicitly addressed our starting point for the project, that traditional exams did not allow all young people to show what they could do and alternative formats could help:

The majority of our learners play video games successfully but struggle to sit and read or write with paper-based activities. These are still capable learners.

Twenty-six replies used words like 'fun', 'interesting', 'enjoyable' or 'engaging' to describe participant responses, one noting that a young person with ADHD had been far more engaged than was usual for them when completing the SBA. A recurring theme was that participants seemed more relaxed and less pressured in the SBA than a traditional exam:

Less formal. Seems less intimidating for students. Reduces the anxiety around exams for students.

A person with ADHD found it more interesting than a plain paper exam.

It was described as seeming 'less like an exam' and 'for those who have anxiety this is less of a challenge' and 'seemed less pressured'. Comments included:

The learner seemed less nervous taking the assessment than if it had been paper based.

Learners are motivated and excited - no anxiety shown.

Although, as the participants' own responses indicate, anxiety was not eliminated for all:

Still certain anxieties regarding how they are progressing/achieving while working through the story. However in reality this would be the same with a more traditional assessment.

As noted, one of our areas of interest was whether the format aided focus and concentration or whether it distracted from the tasks. We did not ask staff a specific question about this, but many addressed it, 22 responses mentioning greater focus, engagement and concentration, often relating it to being more relaxed:

The student was fully focused throughout even though questions set at a much higher level. Attempted every question.

Not all students can sit a paper based exam for hours at a time, they struggle to concentrate and have difficulty processing the question.

Evidence of better focus and engagement with this type of assessment.

2 students said that this helps them focus better.

I have noticed learners focus for longer periods of time when completing these types of assessments. I have also noticed that some students are more relaxed and they confirmed this when we spoke about this afterwards in general conversations.

The students can see where the question is leading and gives them something to keep them hooked until the end.

They seem more relaxed and focused on the task.

Only one staff member made a comment about the format of the SBAs distracting from answering the assessment questions:

It can be very distracting and unclear what is being asked as the story element takes over from the learning.

This was the only staff member who had multiple concerns, commenting:

... they became very frustrated and wound up by it as it wasn't intuitive at all to use... I think with a lot more tweaking it could work but as it stands it does not.

In addition, staff made lots of useful comments about story, images and overall design of the platform, many of which were incorporated into our second set of SBAs. Some also commented on the benefits of having the questions read aloud.

Overall, staff felt their students benefited from a reduction in anxiety and, except for one, did not express concern that the format lessened focus and concentration. In fact, almost all reported the opposite, citing increased focus and concentration as a key benefit of SBAs.

Limitations

There were several limitations to the pilot study. First, participants were well aware that the SBAs were not 'real' high-stakes assessments; the results would not go on their CVs or contribute to their chances of getting an apprenticeship. They were therefore less pressured – though it is worth remembering that the very process of exposing your ability level to judgement can always be challenging, particularly for those with a history of past, and perhaps humiliating, failures. Because of this, students' reports of reduced anxiety and increased confidence need treating with caution. On the other hand, this reduction in pressure can work both ways and demotivate young people. Desire to achieve well in a 'real' assessment can act as an incentive to persevere through boredom or resist the opportunity to go and play football with your friends. It is acknowledged that "fake" exam environments impact the validity of results in any educational study and that this issue is not something unique to our project. As one participant noted:

The sense of importance can be good motivation but also have the opposite effect more often than not.

Second, the year 1 data was less complete than we had planned as the process of exam scrutiny and review through NCFE and the work of the platform developers took longer than expected. This meant that the maths assessment was not ready until July and the English assessment not until the end of the academic year. This meant that only four of the five settings completed the assessments in year 1; we do not have data for the 6th form setting for either assessment for Year 1.

Third, we were reliant on setting staff providing us with learners that were working at the level of Functional Skills Level 1, though in year 1 were given a significant number of learners not yet working at this level. While many of them enjoyed being part of the research and attempted the questions with impressive thought and perseverance, it did mean they could not engage with the content of the questions as planned. We attempted to remedy this in year 2 by reiterating the importance of this to setting staff, and although we were still given a number of learners not yet working at Level 1, the level learners were working at was much more as we had planned it to be in year 2.

Fourth, this was a relatively small study on a group of students all of whom had diagnosed SEND and many of whom were vulnerable more generally. Although we believe SBAs might be beneficial to many students, our study cannot be generalised. Finally, our sample was determined largely by convenience, and although the gender imbalance was in part a reflection of the SEND diagnoses and settings, the almost exclusively white British nature of our sample should be made more representative in future studies.

Summary of results

Most of the young people that sat our assessments did not go on to sit a formal Maths or English exam. They were either deemed not ready to be entered or did not engage with the exam process. The fact that these same learners were willing to engage with our assessment,

indicates a level of comfort with the environment created by our assessments and confidence to be able to complete it. This may be a result of the assessment format making them feel less anxious, or it may be that they knew that this was not high stakes and therefore, the perceived pressure and threat attached was lower. Some learners viewed their involvement in the project as advisors who were offering valuable feedback on the platform and its content. This was incredibly beneficial to us as researchers, looking to consistently make improvements to the quality of the product. However, this did reduce the validity of the data as we are unable to identify to what extent learners attempted to answer the questions as we could not recreate a high stakes environment.

Year 1 data:

The pass rate for the NCFE version of the English paper was (17 marks) 57%. All of the questions in our English paper were replicas, adapted from the original NCFE paper to fit our storyline. 64% of those that sat our English assessment passed with a score equivalent or above the pass rate for the paper. The highest score achieved on our English paper was 77%. All English results are included in the appendix.

The pass rate for the NCFE version of the Maths paper was (36 marks) 60%. All of the questions in our maths paper were replicas, adapted from the original NCFE paper to fit our storyline. None (0%) of those that sat our maths assessment passed with a score equivalent or above the pass rate for the paper. The highest score achieved on our maths paper was 28.9%. All maths results are included in the appendix.

Year 2 data:

The pass rate for the NCFE version of the English paper was (17 marks) 57%. All of the questions in our English paper were replicas, adapted from the original NCFE paper to fit our storyline. 52.2% of those that sat our English assessment passed with a score equivalent or above the pass rate for the paper. The highest score achieved on our English paper was 86.7% but we also had another learner who scored 80% and 10 other learners who achieved scores between 57% and 76%. All English results are included in the appendix.

The pass rate for the NCFE version of the Maths paper was (38 marks) 63%. All of the questions in our maths paper were replicas, adapted from the original NCFE paper to fit our storyline. 6.5% of those that sat our maths assessment passed with a score equivalent or above the pass rate for the paper. The highest score achieved on our maths paper was 84.4% but we also had another learner who scored 75.5%. All maths results are included in the appendix.

The assessments were not 'high stakes' and we could not recreate an authentic high stakes experience offered by a real exam. Learners were aware that their results were not going to be used as a judgement of their ability and although exam style conditions were adhered to, they were undertaken in small rooms with small numbers of other learners and no invigilators. This was positive in terms of their reduced anxiety but negative in terms of the validity of the data obtained.

Conclusion

This project was motivated by our experiences working with young people with SEND, at risk of becoming NEET, have few or no qualifications, and are struggling to gain those they need in order to forge a good future for themselves. For many of those we work with at Really NEET, the current format of examinations has failed them; paper-based exams generate anxiety, frustration and boredom, lessening their concentration, confidence and perseverance. We speculated that an alternative form of assessment, less loaded with past failure and more in tune with their current interests, might meet their needs better. We therefore set about developing such an assessment.

Based on their interest in gaming and computers, and taking into account their criticisms of traditional exams, we devised pilot alternative 'mirror' assessments incorporating Level 1 Functional Skills questions as a potential alternative to the conventional Functional Skills assessments. These were story-based, digital and interactive. We piloted them with a group of young people across different settings, all of whom had a diagnosis of ASC, SEMH or ADHD, or more than one of these. We then asked for their responses to these new assessments through surveys, alongside their staff responses and researcher observations. Our primary aim was to see if the results provided a basis to argue for a larger scale project in which SBAs were piloted with a group of young people as their 'real' Functional Skills assessments and the results compared with a group taking the traditional paper-based exam.

The findings of this pilot project give strong support to the idea that functional skills SBAs could help a substantial number of young people with ASC, SEMH and ADHD. Almost three quarters of our participants felt that SBAs would be of benefit to them and most preferred them to traditional exams. In particular, many said that this was because they felt less anxious and more relaxed in SBAs. Although we are well aware that this could in part be due to them not sitting the SBAs as 'real' assessments, where the results would have had implications for their future, we do not believe this to be the whole story. Their survey responses showed that, for some, the change of format enabled them to move beyond what had too often been a long and painful history of exam failure. In part because of this reduction in anxiety, many students reported increased engagement, ability to focus and concentrate on the tasks. Both the reduction in anxiety and the increase in focus and perseverance also emerged strongly from the staff feedback and researcher observations. Importantly, many of the participants also felt that they would be more likely to attempt future qualifications if assessments were offered as SBAs.

The average time taken to complete our English and Maths papers are shown in the table below and suggest that online assessments do not take as long as traditional exams. This may offer another positive argument for using this exam format over traditional exams for some learners.

Paper	Average time taken to complete Online Story-based paper	Time allocated by NCFE to complete traditional paper
Maths Paper Year 1	52 minutes	90 minutes
Maths Paper Year 2	52 minutes	90 minutes
English Paper Year 1	57 minutes	60 minutes
English Paper Year 2	57 minutes	60 minutes

References

- Brooks, F., E. Klemra, K. Chester, J. Magnusson, and N. Spencer. 2020. *Health Behaviour in School-Aged Children (HBSC): World Health Organization Collaborative Cross National Study (Findings from the 2018 HBSC Study for England)*. England: Centre for Research in Public Health and Community Care.
- Buchanan, Denise, Eleanore Hargreaves, and Laura Quick. 2020. "My Life Is like a Massive Jigsaw with Pieces Missing". How "Lower-Attaining" Children Experience School in Terms of Their Well-Being'. *Education 3-13* 49(8):1–13. doi: 10.1080/03004279.2020.1818269.
- Burke, Emma Gilman, Alexa Hryniuk, and Anna Edmondson. 2020. 'Influence of Test Anxiety on Academic Performance in the Anatomical Sciences'. *The FASEB Journal* 34(S1):1–1. doi: 10.1096/fasebj.2020.34.s1.05608.
- Cassady, Jerrell C., and Ronald E. Johnson. 2002. 'Cognitive Test Anxiety and Academic Performance'. *Contemporary Educational Psychology* 27(2):270–95. doi: 10.1006/ceps.2001.1094.
- Children's Commissioner for England. 2019. *Briefing: The Children Leaving School with Nothing*. Children's Commissioner for England.
- Clarke, Samantha Jane, Daryl J. Peel, Sylvester Arnab, Luca Morini, Helen Keegan, and Oliver Wood. 2017. 'EscapED: A Framework for Creating Educational Escape Rooms and Interactive Games to For Higher/Further Education.' *International Journal of Serious Games* 4(3). doi: 10.17083/ijsg.v4i3.180.
- Coles, Bob, Christine Godfrey, Antonia Keung, Steven Parrot, and Jonathan Bradshaw. 2010. *Estimating the Life-Time Cost of NEET: 16-18 Year Olds Not in Education, Employment or Training Research Undertaken for the Audit Commission*. Department of Social Policy and Social Work and Department of Health Sciences, University of York.
- DfE. 2023. *Special Educational Needs and Disability: An Analysis and Summary of Data Sources*. London: Department for Education.
- van Gaalen, A. E. J., J. Brouwer, J. Schönrock-Adema, T. Bouwkamp-Timmer, A. D. C. Jaarsma, and J. R. Georgiadis. 2021. 'Gamification of Health Professions Education: A Systematic Review'. *Advances in Health Sciences Education* 26(2):683–711. doi: 10.1007/s10459-020-10000-3.
- Gladwell, D., G. Popli, and A. Tsuchiya. 2022. 'Predictors of Becoming Not in Education, Employment or Training: A Dynamic Comparison of the Direct and Indirect Determinants'. *Journal of the Royal Statistical Society: Series A*.
- Hickey, Daniel T., Adam A. Ingram-Goble, and Ellen M. Jameson. 2009. 'Designing Assessments and Assessing Designs in Virtual Educational Environments'. *Journal of Science Education and Technology* 18(2):187–208. doi: 10.1007/s10956-008-9143-1.

- Jackson, Sarah E., Jamie Brown, Lion Shahab, Ann McNeill, Marcus R. Munafò, and Leonie Brose. 2023. 'Trends in Psychological Distress Among Adults in England, 2020-2022'. *JAMA Network Open* 6(7):e2321959. doi: 10.1001/jamanetworkopen.2023.21959.
- Jiménez, Cristina, Nuria Arís, Ángel Magreñán Ruiz, and Lara Orcos. 2020. 'Digital Escape Room, Using Genial.Ly and A Breakout to Learn Algebra at Secondary Education Level in Spain'. *Education Sciences* 10(10):271. doi: 10.3390/educsci10100271.
- Keddie, Amanda. 2016. 'Children of the Market: Performativity, Neoliberal Responsibilisation and the Construction of Student Identities'. *Oxford Review of Education* 42(1):108–22. doi: 10.1080/03054985.2016.1142865.
- Molina-Torres, Guadalupe, Irene Sandoval-Hernández, Carmen Roperro-Padilla, Miguel Rodriguez-Arrastia, Jesús Martínez-Cal, and Manuel Gonzalez-Sanchez. 2021. 'Escape Room vs. Traditional Assessment in Physiotherapy Students' Anxiety, Stress and Gaming Experience: A Comparative Study'. *International Journal of Environmental Research and Public Health* 18(23):12778. doi: 10.3390/ijerph182312778.
- Nadeau, Kathleen. n.d. 'How Can I Study so Hard and Still Do Badly on Tests? – The Chesapeake Center'. Retrieved 21 June 2023 (<https://chesapeakeadd.com/home/education-and-training/articles/how-can-i-study-so-hard-and-still-do-badly-on-tests/>).
- Nicholson, Scott. 2016. 'Ask Why: Creating a Better Player Experience through Environmental Storytelling and Consistency in Escape Room Design'. Lansing, Michigan.
- O'Shea, Nick. 2021. *Covid-19 and the Nation's Mental Health*. London: Centre for Mental Health.
- Park, Namkee, Kwan Min Lee, Seung-A. Annie Jin, and Sukhee Kang. 2010. 'Effects of Pre-Game Stories on Feelings of Presence and Evaluation of Computer Games'. *International Journal of Human-Computer Studies* 68(11):822–33. doi: 10.1016/j.ijhcs.2010.07.002.
- Pearson. 2022. 'Functional Skills Pass Rates 2021/22'. Retrieved 27 June 2023 (<https://www.pearson.com/uk/educators/training-providers/insight-blog/2022/08/functional-skills-pass-rates-2021-22.html>).
- Powell, Andy. 2021. *NEET: Young People Not in Education, Employment or Training*. House of Commons Library.
- Quick, Laura. 2023. 'Making Sense of Low Attainment: Children's Experiences in the Primary Classroom'. Doctoral, UCL (University College London).
- Reay, Diane, and Dylan Wiliam. 1999. "'I'll Be a Nothing": Structure, Agency and the Construction of Identity through Assessment'. *British Educational Research Journal; London* 25(3):343–54.

- Stöber, Joachim, and Reinhard Pekrun. 2004. 'Advances in Test Anxiety Research'. *Anxiety, Stress & Coping* 17(3):205–11. doi: 10.1080/1061580412331303225.
- Terlouw, Gijs, Derek Kuipers, Job Van 'T Veer, Jelle T. Prins, and Jean Pierre E. N. Pierie. 2021. 'The Development of an Escape Room–Based Serious Game to Trigger Social Interaction and Communication Between High-Functioning Children With Autism and Their Peers: Iterative Design Approach'. *JMIR Serious Games* 9(1):e19765. doi: 10.2196/19765.
- Timmis, Sue, Patricia Broadfoot, Rosamund Sutherland, and Alison Oldfield. 2016. 'Rethinking Assessment in a Digital Age: Opportunities, Challenges and Risks'. *British Educational Research Journal* 42(3):454–76. doi: 10.1002/berj.3215.
- Veldkamp, Alice, Joke Daemen, Stijn Teekens, Stefan Koelewijn, Marie-Christine P. J. Knippels, and Wouter R. Joolingen. 2020. 'Escape Boxes: Bringing Escape Room Experience into the Classroom'. *British Journal of Educational Technology* 51(4):1220–39. doi: 10.1111/bjet.12935.
- Zainuddin, Zamzami, Muhammad Shujahat, Hussein Haruna, and Samuel Kai Wah Chu. 2020. 'The Role of Gamified E-Quizzes on Student Learning and Engagement: An Interactive Gamification Solution for a Formative Assessment System'. *Computers & Education* 145:103729. doi: 10.1016/j.compedu.2019.103729.

Appendix A: year 1 data

Feedback post exam (English)					English Score (Paper P001233)	Maths Score (Paper P001255)	Time taken English	Time taken Maths
Project groupings	Images distract?	Confidence?	Preference?	Likelihood to undertake qual in format?	Pass = 17	Pass = 36		
SEMH	Not really	More Confident	OSB	More Likely	10 (33%)	DNS	28 mins	44 mins
SEMH	Not at all	More Confident	OSB	More Likely	17 (57%)	DNS	52 mins	65 mins
SEMH	Not at all	No difference	OSB	More Likely	23 (77%)	6 (13.3%)	48 mins	53 mins
ADHD, ASC & SEMH						5 (11.1%)		35 mins
ADHD, ASC & SEMH						7 (15.5%)		26 mins
ADHD & SEMH								
SEMH	Not at all	More Confident	OSB	More Likely	9 (30%)	DNS	85 mins	60 mins
SEMH	Yes a Little	More Confident	OSB	More Likely	17 (57%)	DNS	75 mins	86 mins
SEMH	Not at all	More Confident	OSB	More Likely	15 (50%)	DNS	72 mins	77 mins
ASC	Not at all	No difference	OSB	More Likely	17 (57%)	1 (2.2%)	83 mins	64 mins
ASC & SEMH						RESULT DIDN'T SAVE		51 mins
ASC & SEMH								
SEMH	Not at all	More Confident	OSB	More Likely	18 (60%)	0 (0%)	64 mins	40 mins
ADHD & SEMH	Neither	More Confident	OSB	More Likely	14 (47%)	13 (28.9%)	36 mins	50 mins
ADHD & SEMH	Not really	More Confident	OSB	More Likely	18 (60%)	4 (8.8%)	31 mins	35 mins
SEMH	Yes a little	More Confident	OSB	More Likely	20 (67%)	10 (22.2%)	51 mins	42 mins

Appendix B: year 2 data

Project groupings	TRADITIONAL VERSION OF EXAMS		ONLINE SBA VERSION OF EXAMS			
	English Score	Maths Score	English Score (Paper P00)	Maths Score (Paper P00)	Time taken English	Time taken Maths
			Pass = 17	Pass = 38	Av = 57	Av = 52
SEMH			DNC	4 / 45 = 8.9%		55
ASC / SEMH			DNC	Score did not save		47
SEMH			7 / 30 = 15.6%	5/45 = 11.1%	61	52
			4 / 30 = 8.9%	2/45 = 4.4%	49	45
ASC			DNC	DNC		
ASC			11 / 30 = 24.4%	8/45 = 17.8%	71	50
ASC / SEMH			DNC	DNC		
SEMH			DNC	1/45 = 2.2%	35	38
SEMH & ASC			DNC			
SEMH			DNC	2/45 = 4.4%		35
SEMH			19/30 = 63.3%	5 / 45 = 11.1%	76	62
SEMH			DNC (Seizures - flashing imagery)	DNC (Seizures - flashing imagery)		
ASC			9/30 = 30%	13 / 45 = 28.8%	50	39
ASC			18/30 = 60%	10 / 45 = 22.2%	60	57
ASC	Failed L1 reading element	Not entered deemed not ready	17/30 = 56.7%	DNC	54	
SEMH	Not entered deemed not ready emotionally		22/30 = 73.3%	38/45 = 84.4%	off-site	off-site
ASC	L2 Pass	EL3 Failed	20 / 30 = 66.7%	DNC	87	
ASC	L2 Pass	Not entered deemed not ready	DNC	DNC		
SEMH	Not entered deemed not ready	Not entered deemed not ready	DNC	DNC		
ASC	L2 Pass	Not entered deemed not ready	23 / 30 = 76.6%	DNC	Not recorded	
ASC	L1 Pass	Not entered deemed not ready	20/30 = 66.7%	19 / 45 = 42.2%	61	55
ASC & SEMH	L2 Failed reading and writing elements	Not entered not deemed ready	DNC	16 / 45 = 35.6%		47
ASC & ADHD	L2 Failed writing element	Not entered not deemed ready	DNC	1 / 45 = 2.2%		55
ASC	Not entered not deemed ready	Not entered not deemed ready	DNC	DNC		
ASC & ADHD	L2 Failed reading and writing elements	Passed EL3	DNC	1 / 45 = 2.2%		40
SEMH	L3 Failed reading and writing elements	Not entered deemed not ready	DNC	5 / 45 = 11.1%		
ASC & SEMH	L2 Failed reading and writing elements	Passed EL3	DNC	3 / 45 = 6.6%		
SEMH	Passed EL3	Passed EL3	DNC	0 / 45 = 0%		
ASC	Not entered not deemed ready	Not entered not deemed ready	DNC	DNC		
ADHD, ASC & SEMH	Passed reading element	Not entered not deemed ready	26 / 30 = 86.7%	26/ 45 = 57.8%	52	67
SEMH	Failed L2	Failed L2	DNC	DNC		
ASC & SEMH	Passed L1	Passed L1	24 / 30 = 80%	34 / 45 = 75.5%	74	90
SEMH	Passed reading element L2	Not entered not deemed ready L2	DNC	DNC		
ASC	Failed L2	Failed L2	DNC	11 / 45 = 24.4%		55
SEMH	EL3 Pass	EL3 Failed	0 / 30 = 0%	DNC	27	
SEMH	Not entered not deemed ready	Not entered not deemed ready	7 / 30 = 23.3%	DNC	34	
SEMH	Passed reading and writing for EL2	Not entered not deemed ready	13 / 30 = 43.3%	7 / 45 = 15.6%	50	59
ASC	Not entered not deemed ready	Not entered not deemed ready	DNC	7 / 45 = 15.5%	Not recorded	Not recorded
SEMH	Not entered not deemed ready	Not entered not deemed ready	DNC	4 / 45 = 8.9%	Not recorded	Not recorded
ASC	Failed writing element L2	Not entered not deemed ready	DNC	0 / 45 = 0%	Not recorded	Not recorded
ASC	Not entered not deemed ready	Not entered not deemed ready	DNC	0 / 45 = 0%	Not recorded	Not recorded
ASC & ADHD						
SEMH			21/30 = 70%	9/45 = 20%	50	52
ASC						
ASC			18/30 = 60%		84	
ASC & ADHD						
ASC						
ASC & ADHD						
ASC			15/30 = 50%	13/45 = 28.9%	50	46
ASC			17/30 = 56.7%	8/45 = 17.8%	58	47
SEMH			DNF			
SEMH			9/30 = 30%	2 / 45 = 4.4%	42	
SEMH			11/30 = 36.7%	16 / 45 = 35.6%	59	52
SEMH			15 / 30 = 50%	17/45 = 37.8%	63	44