

Should open-book, open-web exams replace traditional closed-book exams in Science in higher education? An evaluation of their effectiveness in different disciplines

Laura Roberts
Swansea University, UK

Joanne Berry
Swansea University, UK

Abstract

The mass shift to open-book, open-web (OBOW) assessments during the pandemic highlighted new opportunities in higher education for developing accessible, authentic assessments that can reduce administrative load. Despite a plethora of research emerging on the effectiveness of OBOW assessments within disciplines, few currently evaluate their effectiveness across disciplines where the assessment instrument can vary significantly. This paper aims to evaluate the experience students across different Science subjects had of OBOW exams to contribute to an evidence-base for emerging post-pandemic assessment policies and strategies. In April 2021, following two cycles of OBOW exams, we surveyed students studying different Science degrees across a range of subjects to determine their preparation strategy, experiences during the exam, perception of development of higher order cognitive skills, test anxiety, and how they thought these assessments might enhance employability. Overall, students from subjects that use assessment instruments requiring analytical, quantitative-based answers (Maths, Physics, Computer Science, and Chemistry) adapted their existing study skills less effectively, felt less prepared and experienced higher levels of stress compared to students of subjects using more qualitative discursive based answers (Biosciences and Geography). We conclude with recommendations on how to enhance the use of OBOW exams: these include supporting and developing more effective study skills, ensuring assessments align with intended learning outcomes, addressing the issue of academic integrity, promoting inclusivity, and encouraging authentic assessment. Based on the outcomes of this study, we strongly advise that assessment policies that foster the whole-scale roll-out of OBOW

assessment consider the inter-disciplinary impacts on learner development, staff training, and workload resources.

Keywords: open-book exams; online assessments; STEM; closed-book exams.

Introduction

Covid-19 caused an unprecedented mass global migration to online learning as most Higher Education Institutions (HEI) mandated remote teaching delivery. While teaching online enabled rapid, flexible access, the speed of migration was unprecedented. One major pedagogic shift was the adaptation from traditional, onsite, invigilated closed-book exams (CBEs) to open-book, open-web (OBOW) exams. Universities moved thousands of exams from in-person to online (Rahul, 2020). This approach was common in many Science disciplines as part of an emergency remote teaching strategy.

CBEs have been highly valued across Science as these subjects often require students to have in-depth, rapidly accessible knowledge (Durning et al., 2016). The question style in CBE exams can however differ dramatically between subjects: broadly, quantitative disciplines (Maths, Physics, Chemistry, Computer Science) favour more analytical question instruments with rapid marking producing binary summative outcomes. Disciplines such as Biosciences and Geography tend to use more essay-based questions with discursive outcomes that require high investment in assessing. Many analytical disciplines also rely on invigilated exams as learner authenticity can be more difficult to confirm compared to essay-based instruments (Zhang et al., 2021).

OBOW exams are not new; they are one type of OBE that allows access to external resources (online, course materials) during a set timeframe (Dayananda et al., 2021). They may be completed on or away from campus, usually have a fixed, short duration (two-four hours), and may or may not be invigilated or proctored. All students in a cohort will usually take them within a given timeframe on a specific date, which is what makes them different from Take Home Exams.

There is a significant body of research that compares the merits of traditional CBEs with OBEs (Durning et al., 2016; Johanns, Dinkens and Moore, 2017; Bengtsson 2019; Parker et al., 2019). CBEs and OBEs should not be perceived as different forms of the same assessment, or like-for-like replacements that simply use a different media. CBEs generally ask ‘what do you know?’ while OBEs often ask ‘what can you do with this knowledge?’ Thus, these two assessment methods test very different forms of learning: knowledge vs. intellectual skills. Unfortunately, the emergency response to Covid-19 often conflated these two types of assessment: CBEs became OBEs with often only minimal modifications to the assessment brief.

The mass introduction of OBEs was popular with students (Dayananda et al., 2021) and led to significant savings in administrative services. In addition, OBEs have authentic links to the workplace (Williamson, 2018). These factors have led many to conclude that the mass migration to OBOW assessments was a positive move away from CBEs (Gu et al., 2022).

There are, however, some important drawbacks to OBOW assessments:

1. The emergency switch to online assessments led to higher instances of academic misconduct (Vazquez, Chiang and Sarmiento-Barbieri, 2021). The geographical separation between student and invigilator can intensify ‘riskier’ cheating behaviours, such as collusion and plagiarism (Parker et al., 2019).
2. Implementing an OBOW exam is not as simple as repurposing a closed-book exam (Eurboonyanun et al., 2021; Spiegel and Nivette, 2021; Theophilides and Koutselini, 2000).
3. Teaching strategies need to be developed that help students to prepare for an assessment of this nature (Durning et al., 2016).

Before committing long-term to a new assessment approach, it is important to evaluate whether this approach promotes effective learning and aligns with the discipline pedagogies (Barber et al., 2021). To do this, we designed a survey to evaluate students’ experience of OBOW exams during the pandemic. We wanted to know:

How did students across Science disciplines adapt their study skills to learn for OBOW exams?

There is a correlation between long-term information retention and the effort and depth of pre-exam preparation, yet literature is inconclusive on how students' study for CBEs and OBEs and how this affects performance (Green, Ferrante and Heppard, 2016). Research suggests that students devote more time and effort in studying for invigilated CBEs because they are unable to access external resources (Johanns, Dinkens and Moore, 2017). Conversely, the fact that the cramming and memorisation required for CBEs does not promote conceptual understanding (Myyry and Joutsenvirta, 2015) has led others to argue that students may prepare more effectively for an open book exam because they are required to access and assimilate a range of sources, promoting deeper, more connected learning. Equally, students may also be over-confident and dedicate less effort to studying for an OBE because they can search for answers (Michael, Lyden and Custer, 2019).

Previous research however, fails to consider that different disciplines employ different types of exam question, and this is likely to lead to different approaches in studying behaviour. We currently have little understanding of how students in different Science programmes adapted their study skills to OBOW exams, and if this differed between subjects. Understanding this learning process is essential in ensuring that students are effectively prepared for their assessments.

What did students experience during the exam?

Inequalities in accessibility and affordability of hardware and software, internet connectivity, individual competency with technology, and access to a quiet, unshared space during the time of an exam were all identified as issues affecting students' experience of OBOW exams before the pandemic (Parker et al., 2019). These continued to be a major barrier that prevented access to education to millions across the world during the pandemic (Rahul, 2020). We therefore wanted to assess if our students experienced accessibility issues.

Judging the appropriate duration of an OBOW exam can also be challenging. A balance is needed between allowing enough time to search for and write answers, limiting potential

opportunities for academic misconduct, preventing fatigue, and remaining inclusive. We therefore wanted to know how much of the available time was used by students during the OBOW exams and if this differed between disciplines.

What intellectual skills did students think the OBOW exams helped them to develop?

When students understand what they are required to do in an assessment, they can both adapt their learning as required and recognise how the assessment contributes to their learning development. This encourages engagement and positive emotions (Robinson et al., 2013; Lukasik et al., 2019). With this in mind, if OBOW exams are employed whole-scale, it is critical students understand the skills being developed and their future benefits.

Did students in different disciplines recognise the employability value of the OBOW exams?

Proponents of OBEs argue they are more authentic in relation to real-world practice (Durning et al., 2016). Unstructured problems, requiring the application of knowledge, are provided and students examine an issue from different sources, acting as problem solvers (Parker et al., 2021). This format more closely resembles a realistic work environment where complex issues arise. However, unless students understand this link, they are unlikely to value the experience, and this can lead to disengagement.

Methods

This research was undertaken in the College of Science (CoS), Swansea University, UK. The CoS consists of six departments: Biosciences, Chemistry, Computer Science, Geography, Maths, and Physics, with approximately 2500 students. In March 2020 a rapid shift to online learning took place in reaction to the Covid-19 pandemic. All end-of-year (May/June) exams were switched from in-person to remote OBOW exams. By May 2020, all FHEQ Level 5 and 6 exams were OBOW. The lockdown continued into the next academic year and in January 2021 OBOW exams were used again.

In April 2021 an online questionnaire was developed using Google Forms and deployed via mass email to FHEQ Level 6 students across the six CoS subjects (See Appendix 2). Ethical approval was gained through the CoS Ethics committee (SU-Ethics-Staff-080221/322). Participants were anonymous and completion was voluntary. The survey was split into five main sections: General information, Preparation and Revision, The Exam Experience, Impacts on Learning, and Employability.

The students surveyed had previously experienced in-person, invigilated, closed-book exams in Semester 1 (January) 2019 as FHEQ Level 5 students, meaning they had already developed some of the study skills needed to complete this type of examination. In the following two semesters (Semester 2 May/June 2020 at FHEQ Level 5 and January 2021 of the academic year 2020/21 when they were FHEQ Level 6), these students did OBOW exams in place of the CBEs. All OBOW exams were around two hours in duration. OBOW questions were benchmarked to be answerable within the two-hour time frame, but a submission window of two hours was added to account for potential technical issues during submission. We did not proctor OBOW assessments. Assessment workshops were provided for these students, giving guidance on the study skills required, highlighting the need for critical analysis, synthesis, and problem solving. No past papers were available specifically for these OBOW exams. Assessments were submitted to the Virtual Learning Platform through plagiarism software (Turnitin).

Methods of data analysis

For all multiple-choice answers, differences between analytical and discursive disciplines were analysed in SPSS 21 using a Chi squared test for independence with a critical value of 0.05. To determine if there were differences in the Higher Order Cognitive Skills (HOCS) between analytical and discursive subjects, numerical values were added to the Likert questions (strongly disagree = 1, disagree = 2, neither = 3, Agree = 4, Strongly agree = 5). Chi squared tests for independence were then also used to determine if there were differences between analytical and discursive disciplines in rote and long-term memory and ability to evaluate information and identify relationships between concepts.

For the thematic analysis of the open questions, all responses were reviewed by three reviewers who were trained in standardised reviewing practice. We differentiated responses into specific categories depending on the questions and frequency of occurrence was quantified to give an overall indication of relevance and priority. Questions were designed so that students were not led to perceive that a particular answer was 'correct' and were short to prevent fatigue.

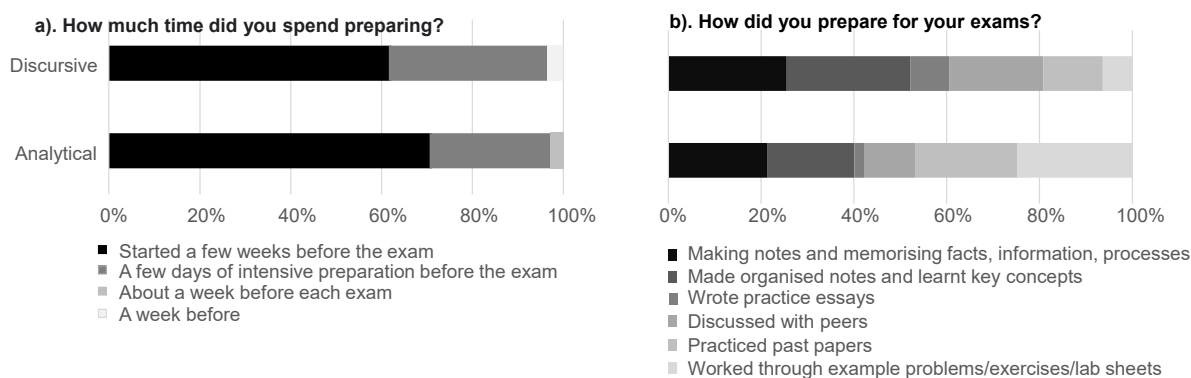
Results

A total of 100 students participated in the survey (27 Biosciences, 23 Chemistry, 28 Computer Sciences, ten Geography, one Maths, and eight Physics students). This amounted to 17%, 74%, 7%, 5%, 1%, and 8% of each cohort respectively. Due to the similarities in OBOW exam question style (quantitative, analytical-based) between the analytical subjects (Maths, Computer Sciences, Chemistry, and Physics) and the more discursive (essay-based) subjects (Biosciences and Geography) and the small sample size, the results have been combined (N = 60 for analytical, N = 37 for discursive).

Preparation and revision

Preparation and revision time for OBOW exams was similar between disciplines with most students starting the revision a few weeks before the exam ($\chi^2_{0.05,3} = 7.15$, $P = 0.21$, Figure 1(a)). Students from analytical and discursive subjects prepared for the exams significantly differently ($\chi^2_{0.05,5} = 25.8$, $P > 0.001$, Figure 1(b)). Most students made notes and memorised information regardless of discipline (65%). 65% of respondents from discursive disciplines selected organising notes and learning key concepts, compared to only 56% of analytical disciplines. Analytical subjects practised past CBE papers and worked through example questions more than discursive subjects (67% vs. 32% and 75% vs. 16% respectively).

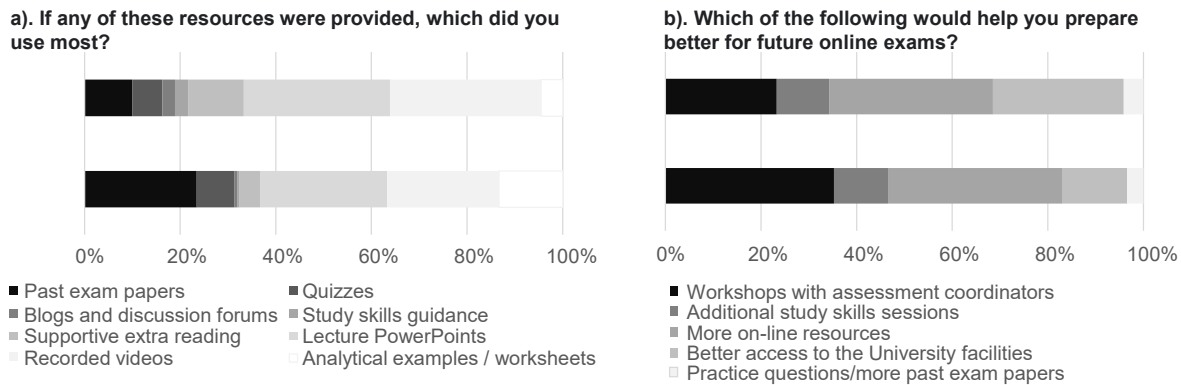
Figure 1. Proportional distribution of survey responses comparing Science students from analytical (N= 60) and discursive (N = 37) subjects relating to (a) how much time was spent preparing (P = 0.21) and (b) how students prepared (P < 0.001) for open-book, open-web exams.



Students in analytical disciplines also used the preparatory resources for OBOW exams differently ($X^2_{0.05,7} = 23.3$, $P < 0.001$, Figure 2(a)); both used lecture recordings (83% analytical, 91% discursive) and recorded videos (73% analytical, 94% discursive) more than any other revision tool. Discursive disciplines used more supportive extra reading (35% compared to 13% for analytical disciplines), while analytical disciplines used past papers (73% compared to 30%) and analytical worksheets (42% compared to 13%).

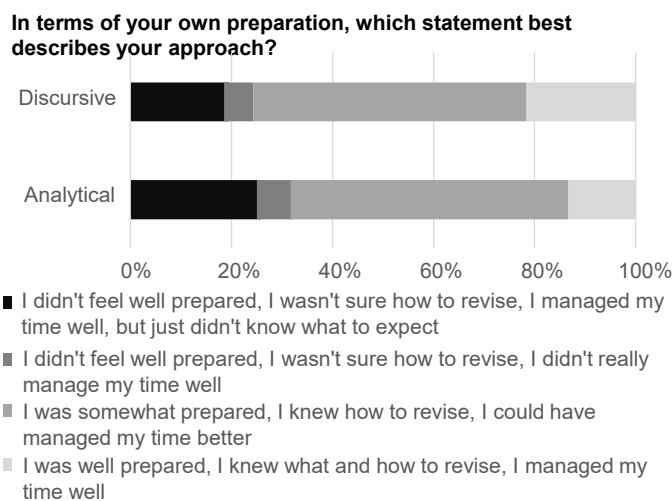
Significant differences were found between discursive and analytical disciplines when students were asked what resources they would need in future to assist their revision for OBOW exams ($X^2_{0.005,5} = 13.28$, $P = 0.01$, Figure 2(b)). Analytical subjects wanted workshops with assessment coordinators (52%), whereas discursive disciplines required access to university facilities (54%).

Figure 2. Proportional distribution of survey responses comparing Science students from analytical (N= 60) and discursive (N = 37) subjects relating to (a) the resources that were used most during studying (P < 0.001) and (b) the resources that would be required (P = 0.001) in future open-book, open-web exams.



No difference was found in how prepared students felt for the OBOW exams ($X^2_{0.05,3} = 1.39$, $P = 0.75$, Figure 3), with most respondents saying that they were somewhat prepared. However, 25% of all students felt that they were not well-prepared and were not sure how to revise; they managed their time well, but just didn't know what to expect.

Figure 3. Proportional distribution of survey responses comparing Science students from analytical (N= 60) and discursive (N = 37) subjects relating to how prepared they felt for OBOW exams (P = 0.75).



When asked to describe how their learning was different to traditional CBEs, 22 different themes were identified, ten were considered positive and 12 were considered negative (Table 1 (Appendix 1)). Significant differences were found between subject types with analytical subjects identifying 15 negative indicators, while discursive subjects identified six negative indicators ($X^2_{0.05,1} = 4.98$, $P = 0.021$, Table 1(a)). The main negatives were did not learn, learnt less, and lost learning faster; the positives were better understanding of concepts and a less stressful environment.

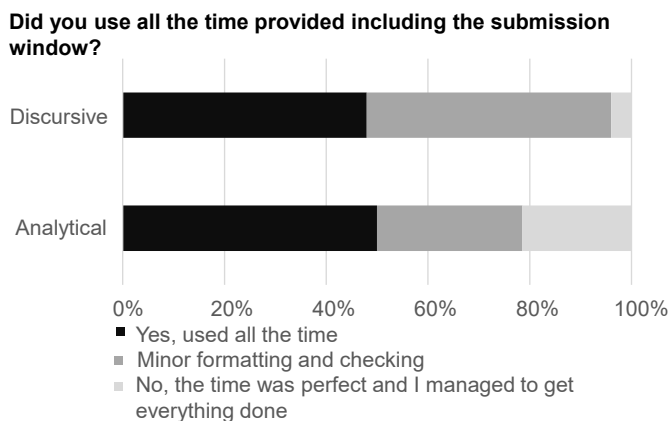
Students were asked to summarise how their revision strategy differed from CBEs: 29 different study skills were identified (Table 1(b)), but no difference was found between analytical and discursive disciplines ($X^2_{0.05,1} = 1.8$, $P = 0.18$). The main positives ($N = 17$) were less memorisation and greater understanding of concepts and organising notes, although some students saw no differences (13%) and some thought that the OBOW exams devalued their degree (less rigour), led to more work, and caused more panic of the unknown.

During the revision period students experienced several challenges with remote learning with 62% finding it difficult to find a quiet space, 51% with internet connection issues, 36% with issues accessing teaching resources, and 29% with PC issues.

The exam experience

Overall, 78% of students managed to complete the exam in time and there was no difference between disciplines ($X^2_{0.05,1} = 0.002262$, $P = 0.98$). However, when asked if the submission window was used for completing the exam, 21% of analytical subjects responded with 'no, the time was perfect', compared with only 4% for discursive subjects (Figure 4). This difference between subjects was not significant ($X^2_{0.05,1} = 4.85$, $P = 0.06$), but it does suggest students studying discursive subjects needed more time than was allocated. Only 44% of students did not experience any technical issues during the exam: 22% had internet failures, 21% had submission issues, 19% had errors in the script, and 12% had PC issues and uploading issues.

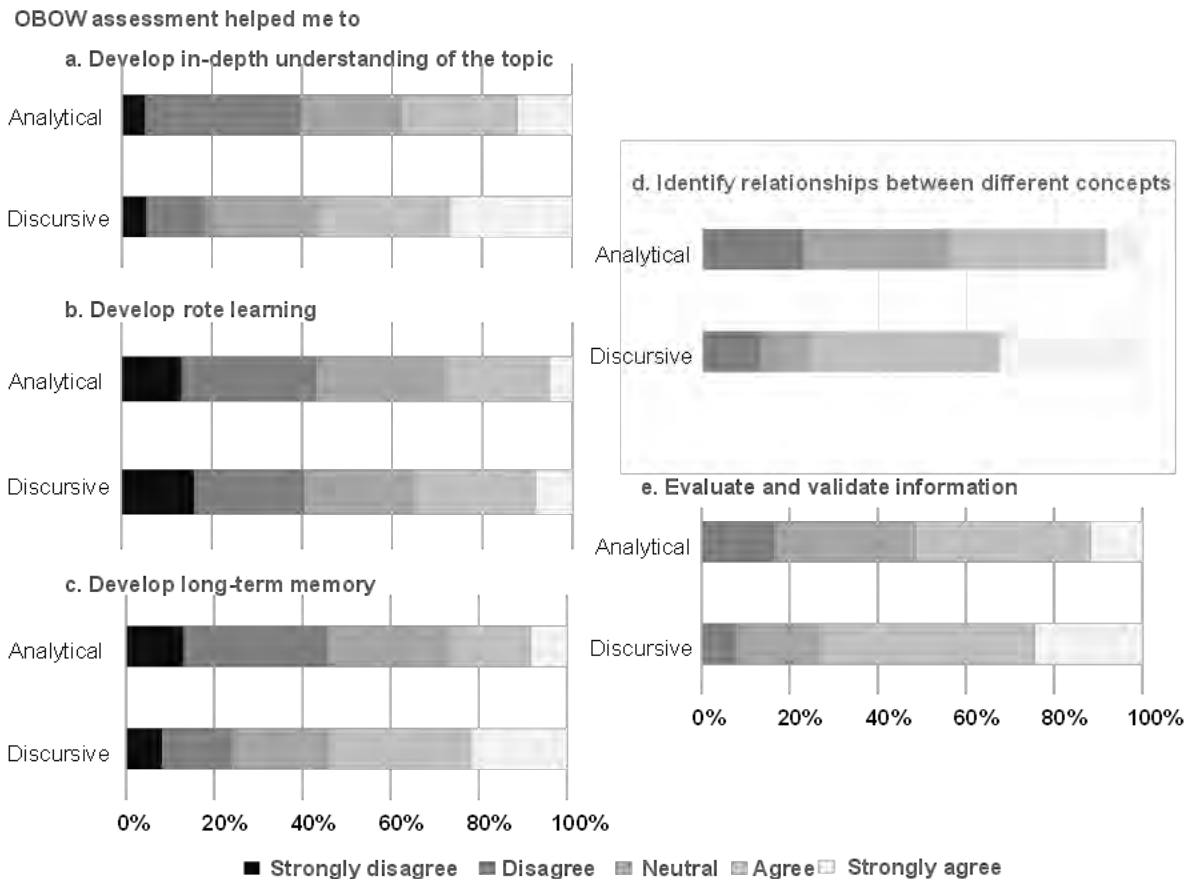
Figure 4. Proportional distribution of survey responses comparing Science students from analytical (N= 60) and discursive (N = 37) subjects concerning how long it took them to complete the OBOW exam (P = 0.06).



Learning and intellectual skills

When comparing the impacts on learning between analytical and discursive disciplines, no difference was observed between how students felt the OBOW assessment helped them to develop in-depth understanding of the topic ($X^2_{0.05,4} = 6.76$, $P = 0.15$), with rote learning and memorisation ($X^2_{0.05,4} = 1.02$, $P = 0.75$) or long-term memory of the topic ($X^2_{0.04,1} = 8.41$, $P = 0.06$), or the ability to evaluate and validate information ($X^2_{0.05,3} = 5.26$, $P = 0.35$). A significant difference was found when students were asked if OBOW exams helped them to identify relationships between different concepts ($X^2_{0.05,3} = 13.81$, $P = 0.01$) (Figure 5), with more students from discursive disciplines strongly agreeing with this.

Figure 5. Analytical (N = 60) and discursive (N = 37) subjects questionnaire responses comparing open-book, open-web exams developed (a) in-depth understanding, (b) rote learning, (c) long-term memory, (d) relationships between concepts, and (e) the ability to evaluate and validate information.



Employability

An open question asking what impacts the exams had on employment revealed 12 positive themes, and two negatives (Table 1(c)). 17 students made negative comments; the main comment (16 responses) was that OBOW had not improved their employability. 34 students made positive comments: ten students identified the real-world applications of the assessment, and ten thought they could identify and prioritise/evaluate information more quickly. No significant difference was found between analytical and discursive disciplines ($\chi^2_{0.05, 1} = 1.9, P = 0.25$).

Test anxiety

When asked if the OBOW exams were stressful, 60% of students from analytical disciplines found them less, or somewhat less stressful than CBEs, 17% found them the same, and 27% found them more stressful. For discursive disciplines however 79% found them less, or somewhat less stressful than CBEs, 26% the same and only 8% found them more stressful. Test anxiety did not differ significantly between subject areas however ($\chi^2_{0.05,4} = 7.4$, $P = 0.10$).

When students were asked if there was anything else they would like to comment on, 36 responded, providing eight positive themes and ten negative themes (Table 1(d)). Analytical subjects had more dissatisfaction, highlighting a disconnect between teaching and assessment, a bad experience overall, and more stress due to no past papers, whereas positive comments referred more often to the real-world applications and less stressful environment ($\chi^2_{0.05,1} = 4.59$, $P = 0.04$).

Discussion

The study shows that though the OBOW exams were similar in duration and support provided, students from different disciplines prepared differently and used distinct support material. There were also differences in overall satisfaction and their perception of the efficacy of their preparation. All students felt that HOCS skills were developed by OBOW exams. Most students, regardless of subject, thought that the OBOW exams positively contributed to enhancing their employability. Many students however experienced issues with remote digital access and administrative errors.

Preparation and revision

Many studies comparing CBEs and OBEs have focused on the duration of study time (Durning et al., 2016; Myyry and Joutsenvirta, 2015). However, few have focused on the actual study skills used by students. For both analytical and discursive subjects, we found the predominant form of revision strategy was to make notes and memorise information. In OBOW exams, the need to memorise facts is almost redundant, as well-organised notes of key concepts and access to material should allow students to formulate an accurate

response (Bengtsson, 2019). It appears our students habitually returned to practising rote memorisation. Recalling facts during a time-restricted assessment is cognitively demanding and may have removed time for more complex thought processes and critical thinking, leading to some dissatisfaction and lack of preparation felt by students (Spiegel and Nivette, 2021). Though many students did adapt their preparation, our findings suggest students would still benefit from greater opportunities to practise appropriate study techniques. Our students had received the results of their assessment prior to taking this questionnaire, which could have perhaps skewed their memory towards what they felt they should have done rather than what they did during revision.

Analytical subjects were more dissatisfied with their preparation and learning experience with the OBOW exams, with some claiming that they were harder, less rigorous, and learning was lost faster. This was probably linked to the type of questions asked and the extent to which they needed to be adapted for the online format. OBOW discursive essays generally differ from CBEs in what students are required to 'do' with the information (for example, critically analyse as opposed to describe). The format otherwise is very similar, and students are generally well-practised at essay writing. For analytical disciplines, more radical changes to the assessment are often needed to negate academic misconduct. In a CBE, a student is assessed on producing the correct numerical answer to a problem. In an equivalent OBOW exam students may be asked to explain the processes, interpret the problem, or add some level of application to the answer (Spiegel and Nivette, 2021). Our results suggest that the traditional approach that analytical students take to revision, such as using worksheets and past papers, may not have effectively developed these skills, and students themselves indicated teaching did not prepare them for this more complex type of examination that diverges from discipline norms.

OBOW assessment duration

Most of the assessments considered here consisted of a two-hour OBOW exam, with an additional two-hour submission window. A quarter of students felt the time was not long enough and most used the exam and submission window to work even though questions were answerable within the first two hours. This was particularly the case in essay-based subjects. Searching and synthesising information and assimilating an answer takes more time than producing information from memory. This confirms previous research: Myyry and

Joutsenvirta (2015) found only 5% of students used less time to answer an OBOW exam question, and over half of the students (55%) reported they used more time to respond when compared to a CBE.

Impacts of OBOW on Higher Order Cognitive Skills (HOCS)

Our students felt that the OBOW exams had developed their HOCS (identifying relationships between concepts and evaluating and validating information) more than rote learning and memorisation. This aligns with many other studies that have found that OBEs develop HOCS (Heijne-Penninga et al., 2011; Malone et al., 2021), but we recognise that our results are limited in that they are based on a single open cohort study and the perception of the students surveyed. We did not directly test for enhancements in these skills or knowledge by using randomised control trials on learning achievement or compare longitudinal data. Students 'opinions' may well have been influenced by information provided by the faculty during the preparation time. Indeed, much of the research into how OBOW exams develop HOCS has been based on qualitative research findings (Johanns, Dinkens and Moore, 2017). More high-quality empirical research is required to ascertain if OBEs are better than CBEs at developing HOCS.

Employability and real-world applications of OBOW exams

One major limitation for CBEs is their separation from real-world contexts. OBOW assessments more accurately simulate real working conditions and offer the opportunity to employ a diverse range of authentic instruments (Williams and Wong, 2009). Being able to locate information competently and accurately is an essential modern life skill. At the very least, employers expect basic competencies in digital skills, including mastering basic computer applications and internet searching (Stowell, 2015). Assessments that emulate competencies needed in real-life situations can motivate students to adopt deep learning strategies and improve generic skills, like problem-solving and written communication (Gulikers et al., 2006).

20% of respondents to our study recognised OBOW exams real-world applications (similar results were obtained by Myyry and Joutsenvirta, 2015). However, approximately 20%

either thought the OBOW exams did nothing for their employability or devalued the degree. This suggests two things:

- i. Students would benefit from additional support during preparation, so they are able to identify the real-world applications.
- ii. Staff require guidance on how to create teaching activities and questions that align with professional competencies.

Embracing these factors will be critical if OBOW exams are to be more widely adopted in the HE assessment toolkit.

Test anxiety and overall experience of OBOW exams

The high-stakes nature of traditional CBEs can cause test anxiety (physiological arousal, emotionality, procrastination, avoidance) in students, weakening academic performance (Tyng et al., 2017). However, although negative emotions were once thought to have exclusively deleterious performance effects, contemporary theories of emotion suggest this assumption is oversimplified (Robinson et al., 2013; Lukasik et al., 2019). Negative emotions like anxiety might motivate students to study more for CBEs, which could result in a superior performance when compared with an unstressed student preparing for OBEs (Myry and Joutsenvirta, 2015). Nevertheless, reducing exam anxiety is often reported to be a significant motivation for employing OBEs (Goothy et al., 2019). Having access to notes, textbooks, and the internet allows cognitive offloading that provides a reprieve from the intense environment of a CBE (Morrison and Richmond, 2020). However, other aspects of OBEs may be anxiety-provoking for students (Gharib, Phillips and Mathew, 2012), including the belief that examiners will set harder questions (Stowell, 2015).

27% of students in analytical disciplines found the OBOW exams more stressful than CBEs, compared to 8% of discursive subjects, mirroring the students' overall experience with the OBOW exams. It is important to highlight here that this was not the view of all students surveyed in the analytical subjects. However, it does suggest that there is more work to be done when using OBOW assessments in analytical disciplines.

Digital access

Students reported many issues with the remote delivery of OBOW exams throughout the pandemic. Lack of a quiet space, internet connection issues, issues accessing teaching resources, and PC issues all hampered students. During the OBOW exam window, nearly a quarter of responders had internet failures, submission issues, or errors in the script, and 12% had PC issues and problems uploading their exams. UK nationwide surveys found very similar results (Barber et al., 2021). A report in 2020 (Maguire, Dale, and Pauli, 2020) recommended that HEIs strengthen their response to digital poverty; this remains a priority concern and additional resources are needed to ensure that all students can access learning materials online.

Further research

Though not captured in the present data, despite efforts to minimise academic misconduct (reduced assessment window, individual interpretation components, randomised assignment allocation), we recorded significant increases (collusion, plagiarism, commissioning) across all disciplines in the switch to OBOW assessments. This is in line with recent research (e.g., Ebaid, 2021; Zhang et al., 2021) and is of considerable concerns with the rise of Artificial Intelligence and natural language processing tools. OBEs can mitigate academic misconduct if the questions target HOCS (López and Whittington 2011; Michael, Lyden and Custer, 2019); students should be challenged with evaluating information and formulating individual responses (Michael, Lyden and Custer, 2019). This is a significant challenge in Science disciplines since grading longer answers that more robustly test an individual creates a significant increase in workload. Faculty workloads, particularly in analytical disciplines, are not aligned to permitting more time-consuming marking, and research around workload norms will be required for these assessments to be rolled out whole-scale.

Furthermore, for most students the exam was impacted by the technical process which is a significant issue if this difficulty either impacted the mark awarded or contributed to anxiety at the time, before, or during subsequent exams. Further research should be aimed at addressing the link between technical issues experienced, experience of stress

and assessment outcomes, particularly as these factors are likely to disproportionately impact neurodiverse students requiring reasonable adjustments.

OBOX Guidance in post-pandemic higher education

There is a rich body of research on the benefits of OBOX assessments (Durning et al., 2016; Parker et al., 2021), so we focus here specifically on guidance for instructors and faculties that wish to embed the use of OBOX exams in a post-pandemic framework (Jisc, 2020; 2021). We have identified six main factors that can enhance student learning and experience.

Develop relevant, effective teaching, learning and assessment strategies to prepare students for OBOX assessments

OBOX exams are different from CBEs. They require different skills, ask different questions, and demand different learning strategies (Johanns, Dinkens and Moore, 2017). These differ between disciplines and appropriate subject-specific approaches should be investigated. Teaching strategies should prepare students for the method of assessment (Gu et al., 2022); if OBOX exams are employed, critical thinking, deep learning and HOCS need to be embedded in teaching (Parker et al., 2021; Malone et al., 2021). Instructors should use questions that require students to locate information and actively do something with it (analyse, evaluate, create, etc.). Skills that allow a student to quickly identify and prioritise key information should be targeted for development.

Develop assessment strategies to encourage academic integrity

Assessments need to be robust and authentic. Quantitative, summative analytical questions provide the greatest opportunities for collusion in an un-proctored/non-invigilated exam. Proctoring software currently has some significant limitations (Parker et al., 2021). But it is highly likely the growth of online learning will lead to more sophisticated software being produced (Barber et al., 2021). To address these issues, instructors should try to develop more authentic questions that include an element of interpretation and/or reflection on how the student arrived at the answer.

Train students to manage their time effectively in OBOW assessments

All students require careful guidance on how to allocate their time during the exam window to ensure they spend the correct amount of time searching and reading compared to writing. We suggest that both teachers and students are allowed to formatively practise OBOW assessments and are encouraged to reflect on the efficacy of their time allocation.

Consider the needs of students with learning disabilities

Many students will use the entire time allowed to work on their OBOW assessment. Although they view having this extra time as an advantage, if they have multiple assessments in an assessment period, they may experience fatigue. This can be much worse for students who have specific learning disabilities that affect reading and writing time (Block, 2012). We did not collect information about the number and characteristics of disabled students. Including this information would have allowed a more rigorous investigation into how to support students with specific learning requirements more effectively. More research is required to appropriately address this complex issue, but we recommend that instructors consider alternative adjustments that allow these students to meet the learning outcomes without experiencing fatigue.

Align OBOW assessments with authentic, real-world scenarios

OBOW assessments provide the ideal opportunity to prepare learners for real-world working. Employability has been high on the HEI agenda for three decades and OBOW exams align with many of these principles of more authentic assessment and deliver on many of the key graduate life-skills. It is important that students are made aware of the skills gained and their relevance to their future work. We recommend that this is made explicit to students as they start to prepare for OBOW exams.

Eliminate digital barriers

The predicted growth of online learning and assessment means we need to find strategies to enhance digital access, build digital skills, and harness technology more effectively; at the same time, we need to ensure inclusivity, accessibility, and equality (Maguire, Dale and Paulie, 2020; Barber et al., 2021). These barriers must be considered by instructors.

For example, provide quiet university spaces with good internet access and laptop loan schemes.

Conclusion

The almost universal adoption of OBOW exams during lockdowns has challenged traditional closed-book, unseen, invigilated exams as the dominant form of assessment in Science HE. OBOW exams have many advantages, but they are not without their barriers. Our study shows that not all students experience OBOWs in the same way and that there can be significant differences in study preparation and overall satisfaction between students of different disciplines. Fundamental questions remain around the validity and rigour of OBOW assessment, which can be exacerbated in different Science disciplines. More broadly, additional research is required on the equitable provisions of digital accessibility and the unknown impacts of OBOW assessment on neurodiverse students. Our research is challenged by being in one area of one institution during a time of unstable change and as such is limited on its wider lessons in the longer term. But the research is well supported within the breadth of literature and this study adds value as part of an ongoing conversation regarding examinations as an assessment format. OBOW exams need to be effectively embedded into curriculum design and pedagogic methods must be aligned so that students develop the skills and confidence they need to succeed in OBOW exams. There is also a need to develop and embed strategies to encourage and promote academic integrity; such strategies will challenge cultural norms and require investment in time, technology, and resources. Undoubtedly, going forward, it will be critical to cement digital skills and accessibility to enable all to fully benefit from a whole-scale adoption of this assessment approach.

Acknowledgements

The authors did not use generative AI technologies in the creation of this manuscript.

References

- Barber, M., Bird, L., Fleming, J., Titterington-Giles, E., Edwards, E. and Leylands, C. (2021) *Gravity assist: Propelling higher education towards a brighter future*. Available at: <https://www.officeforstudents.org.uk/publications/gravity-assist-propelling-higher-education-towards-a-brighter-future/> (Accessed: 20 March 2021).
- Bengtsson, L. (2019) 'Take-home exams in higher education: a systematic review', *Education Sciences*, 9(4). <https://doi.org/10.3390/educsci9040267>.
- Block, R. M. (2012) 'A discussion of the effect of open-book and closed-book exams on student achievement in an introductory Statistics course', *PRIMUS* 22(3), pp.228-238. <https://doi.org/10.1080/10511970.2011.565402>.
- Dayananda, R., Patil, M., Manjunath, S.N., Parshuram, R. and Kautilya. V. (2021) 'Study of students' perception regarding open book assessment and closed book exams', *Indian Journal of Forensic Medicine and Toxicology*, 15(1), pp.946-949. <https://doi.org/10.37506/IJFMT.V15I1.13537>.
- Durning, S. J., Dong, T., Ratcliffe, T., Schuwirth, L., Artino, A. R., Boulet, J. R. and Eva, K. (2016) 'Comparing open-book and closed-book examinations', *Academic Medicine*, 91(4). <https://doi.org/10.1097/ACM.0000000000000977>.
- Ebaid, I. E. S. (2021) 'Cheating among Accounting students in online exams during Covid-19 pandemic: exploratory evidence from Saudi Arabia', *Asian Journal of Economics, Finance and Management*, 4, pp.9-19. Available at: <https://globalpresshub.com/index.php/AJEFM/article/view/1068> (Accessed: 21 June 2021).
- Eurboonyanun, C., Wittayapairoch, J., Aphinives, P., Petrusa, E., Gee, D. W. and Phitayakorn, R. (2021) 'Adaptation to open-book online examination during the Covid-19 pandemic', *Journal of Surgical Education*, 78(3), pp.737-739. <https://doi.org/10.1016/j.jsurg.2020.08.046>.

- Gharib, A., Phillips, W. and Mathew, N. (2012) 'Cheat sheet or open-book? A comparison of the effects of exam types on performance, retention, and anxiety', *Journal of Psychology Research*, 2(8). <https://doi.org/10.17265/2159-5542/2012.08.004>.
- Goothy, S. S. K., Suphal, S., Bandaru, T. S., Movva, S., Manyam, R. and Raju, V. R. (2019) 'Comparison of academic performance and stress levels in open book test and closed book test and perceptions of undergraduate dental students', *MOJ Anatomy & Physiology*, 6(2). <https://doi.org/10.15406/mojap.2019.06.00246>.
- Green, S. G., Ferrante, C. J. and Heppard, K. A. (2016) 'Using open-book exams to enhance student learning, performance, and motivation', *The Journal of Effective Teaching*, 16(1), pp.19-35.
- Gu, S., Yuan, W., Zhang, A., Huo, G., Jiang, M., Han, J. and Shen, N. (2022) 'Online re-examination of postgraduate medical students during the Covid-19 pandemic', *BMC Medical Education*, 22(1), 42. <https://doi.org/10.1186/s12909-022-03100-8>.
- Gulikers, J. T. M., Bastiaens, T. J., Kirschner, P. A. and Kester. L. (2006) 'Relations between student perceptions of assessment authenticity, study approaches and learning outcome', *Studies in Educational Evaluation*, 32. <https://doi.org/10.1016/j.stueduc.2006.10.003>.
- Heijne-Penninga, M., Kuks, J. B. M., Hofman, W. H. A. and Cohen-Schotanus, J. (2011) 'Directing students to profound open-book test preparation: the relationship between deep learning and open-book test time', *Medical Teacher*. 33(1), pp.16-21. <https://doi.org/10.3109/0142159X.2011.530315>.
- Jisc. (2020) *Assessment rebooted from 2020's quick fixes to future transformation. From fixes to foresight: Jisc and EmERGE Education insights for universities and startups*. Available at: <https://repository.jisc.ac.uk/7854/1/assessment-rebooted-report.pdf> (Accessed: 6 November 2020).
- Jisc. (2021) *The future of assessment: five principles, five targets for 2025*. Available at: <https://www.jisc.ac.uk/reports/the-future-of-assessment> (Accessed: 8 June 2021).

- Johanns, B., Dinkens, A. and Moore, J. (2017) 'A systematic review comparing open-book and closed-book examinations: evaluating effects on development of critical thinking skill', *Nurse Education in Practice*, 27. <https://doi.org/10.1016/j.nepr.2017.08.018>.
- López, J. and Whittington, M. S. (2001) 'Higher-order thinking in a college course: a case study', *NACTA Journal*, 45(4), pp.22-29.
- Lukasik, K. M., Waris, O., Soveri, A., Lehtonen, M. and Laine, M. (2019) 'The relationship of anxiety and stress with working memory performance in a large non-depressed sample', 10(4), pp.1-9. <https://doi.org/10.3389/fpsyg.2019.00004>.
- Maguire, D., Dale, L. and Pauli, M. (2020) 'Learning and teaching reimagined: a new dawn for higher education? Exploring the 2020 experience as well as the changing aspirations of the nature and shape of learning and teaching for the future'. Available at: <https://repository.jisc.ac.uk/8150/1/learning-and-teaching-reimagined-a-new-dawn-for-higher-education.pdf> (Accessed: 31 August 2023).
- Malone, D. T., Chuang, S., Yuriev, E. and Short, J. L. (2021) 'Effect of changing from closed-book to formulary-allowed examinations', *American Journal of Pharmaceutical Education*, 85(1). <https://doi.org/10.5688/ajpe7990>.
- Michael, K., Lyden, E. and Custer, T. (2019) 'Open-book examinations (OBEs) in an Ultrasound Physics course: a good idea or a bad experiment?', *Journal of Diagnostic Medical Sonography*, 35(3). <https://doi.org/10.1177/8756479318821075>.
- Morrison, A. B. and Richmond, L. L. (2020.) 'Offloading items from memory: individual differences in cognitive offloading in a short-term memory task', *Cognitive Research*, 5(1). <https://doi.org/10.1186/s41235-019-0201-4>.
- Myyry, L. and Joutsenvirta, T. (2015) 'Open-book, open-web online examinations: Developing examination practices to support university students' learning and self-efficacy', *Active Learning in Higher Education*, 16(2). <https://doi.org/10.1177/1469787415574053>.

- Parker, A., Watson, E., Dyke, N. and Carey, J. (2021) 'Traditional versus open-book exams in remote course delivery: a narrative review of the literature' in *Proceedings 2021 Canadian Engineering Education Association (CEEA-ACEG21) Conference*, Charlottetown, PE, Canada 21-23 June. (2021) pp.1-7.
- Rahul, K. (2020) 'Assessing higher education in the Covid-19 era', *Journal of Educational Research and Practice*, 29(2), pp.37-41.
<https://doi.org/10.26522/brocked.v29i2.841>.
- Robinson, O. J., Vytal, K., Cornwell, B. R. and Grillon, C. (2013) 'The impact of anxiety upon cognition: perspectives from human threat of shock studies', 7(203).
<https://doi.org/10.3389/fnhum.2013.00203>.
- Spiegel, T. and Nivette, A. (2021) 'The relative impact of in-class closed-book *versus* take-home open-book examination type on academic performance, student knowledge retention and wellbeing', *Assessment & Evaluation in Higher Education*, 48(1), pp.1-14. <https://doi.org/10.1080/02602938.2021.2016607>.
- Stowell, J. R. (2015) 'Online open-book testing in face-to-face classes', *Scholarship of Teaching and Learning in Psychology*, 1(1). <https://doi.org/10.1037/stl0000014>.
- Theophilides, C. and Koutselini, M. (2000) 'Study behavior in the closed-book and the open-book examination: a comparative analysis', *Educational Research and Evaluation*, 6(4). <https://doi.org/10.1076/edre.6.4.379.6932>.
- Tyng, C. M., Amin, H. U., Saad, M. N. M. and Malik, A. S. (2017) 'The influences of emotion on learning and memory', *Frontiers in Psychology*, 8, 1454.
<https://doi.org/10.3389/fpsyg.2017.01454>.
- Vazquez, J. J., Chiang, E. P. and Sarmiento-Barbieri, I. (2021) 'Can we stay one step ahead of cheaters? A field experiment in proctoring online open book exams', *Journal of Behavioral and Experimental Economics*, 90, 101653.
<https://doi.org/10.1016/J.SOCEC.2020.101653>.

Williams, J. B. and Wong, A. (2009) 'The efficacy of final examinations: a comparative study of closed-book, invigilated exams and open-book, open-web exams', *British Journal of Educational Technology*, 40(2). <https://doi.org/10.1111/j.1467-8535.2008.00929.x>.

Williamson, M. H. (2018) 'Online exams: the need for best practices and overcoming challenges', *The Journal of Public and Professional Sociology*, 10(1).

Zhang, L. Y., Petersen, A. K., Liut, M., Simion, B. and Alaca, F. (2021) 'A multi-course report on the experience of unplanned online exams', in *Proceedings of the 52nd ACM Technical Symposium on Computer Science Education*. Online, USA 13-20 March. New York: Association for Computing Machinery, pp.17-23. <https://doi.org/10.1145/3408877.3432515>.

Appendices

Appendix 1

Table 1. Summary of responses to open questions relating to students experience of OBOW exams. A = analytical subjects, D = discursive subjects.

a. Did learning differ during the revision?					
Positive responses			Negative responses		
	A	D		A	D
Better understanding of concepts	4	4	Did not learn	1	2
Better long-term memory	0	2	Learnt less than a formal exam	1	0
Less formal	0	1	Learning lost faster	1	0
Lecture recordings beneficial	0	1	More general	0	1
Less formal	0	1	Loss of discipline/motivation	3	1
Less stressful	2	1	Peer isolation	0	1
More flexible	0	1	More work	0	1
More independent	0	1	Harder	2	0
Better referencing	1	0	Lecturers disrupted/difficult	2	0
Less memorising	7	4	Home disruptions	1	0
			Poor quality teaching	3	0
			Teaching not adapted to new approach	1	0
Total responses	14	16		15	6

b. How did your revision strategy differ?					
Positives			Negatives		
	A	D		A	D
Less memorising	18	16	No difference	9	4
Understand concepts	9	6	Started later	2	1
Organised notes	4	7	Less extra reading	0	2
More further reading	0	3	Revised less	1	1
Less stress	3	0	No library	1	1
Access lecture recordings	1	1	More work	1	0
More focus on past-papers	1	1	More effort	1	0
Easier	2	0	Less rigour	1	0
Focus on analysis	0	1	Exams de-valued	1	0
Build essay plans	0	1	On-line harder	1	0
Detailed notes on one document	0	1	Did not do past exams	1	0
Quizzes	1	0	More panic of the unknown	1	0
Practice questions	1	0			
Revised both on-line and off-line	1	0			
More problem solving	1	0			
Better being at home	1	0			
Manged time better	1	0			
Total responses	44	37		20	9

c. Impacts on employability?					
Positives			Negatives		
	A	D		A	D
More real world applications	5	5	Nothing	11	5
Write faster essays	0	1	Less rigorous, devalued degree	0	1
Time management	0	2			
Identify and prioritise/evaluate info	6	4			
Read more resources	0	1			
Learn independently	0	1			
Better organisation	1	2			
Able to deal with multiple deadlines	0	2			
Critically analyse information	0	1			
Stress management	1	0			
More remote jobs available	1	0			
Better work/life balance	1	0			
Total	15	19		11	6

d. Anything to add?					
Positives			Negatives		
	A	D		A	D
Prefer relaxed environment	0	2	Bad experience	2	0
Less overloaded	0	1	Devalued degree	1	0
Show more in-depth knowledge	0	1	More stress as fewer past papers	2	0
Fewer mistakes	0	1	Less confident	1	0
Real-world skills	3	0	Harder exams due to less guidance	1	0
Less memory focused	0	1	Disconnect between teaching delivery and on-line learning style	2	1
Way forward	1	0	Less effective	1	0
Helped mental health	1	0	Harder to prepare for	1	0
			Not a university experience	0	1
			Less motivated	1	0
Total responses	5	6		12	2

Appendix 2

Questionnaire produced to assess how learners in Science higher education in Swansea University's College of Science perceived OBOW assessments during the COVID pandemic

CoS Science Assessment Questionnaire

The purpose of this questionnaire is to get an insight into how students have experienced on-line assessments during the COVID-19 pandemic.

This questionnaire is anonymous and personal information will not be publicly accessible.

The aim of this platform is to understand how students have prepared for the on-line assessments. Please refrain from identifying specific lectures and only refer to module codes or titles if required.

The questionnaire should take around 20 minutes to complete. There are four main sections: General information, Preparation and Revision, The Exam and Impacts on Learning. There are 27 question.

Please make every effort to complete the questions as honestly and thoroughly as possible.

This questionnaire is completely voluntary and you are free to leave out any questions or withdraw from the survey at any time.

We thank you in advance for the time taken to do this.

General information

Please complete this set of general demographic questions

1. What department/college are you from?

Mark only one oval.

- Computer Science
- Maths
- Physics
- Chemistry
- Geography
- Biosciences
- Other: _____

2. What year of study are you in?

Mark only one oval.

- Foundation (Level 3)
- Year 1 (Level 4)
- Year 2 (Level 5)
- Year 3 (Level 6)
- Intercalary Year (i.e. Year Abroad, Year in Industry)
- PGT (Level 7: Masters)
- Other: _____

3. Please select your age range

Mark only one oval.

- 18 - 23
- 24 - 30
- 31 - 50
- 51 +

4. Are you:

Mark only one oval.

- Home (UK National)
- Home (European)
- International (outside Europe)

5. If international, please state home country

On-line exam preparation

This section is aimed at determining how students prepared for their exams.

6. If any of these resources were provided, which did you use most? (please note, not all of these may be relevant to your discipline)

Check all that apply.

- Past exam papers
- Quizzes
- Blogs and discussion forums
- Study skills guidance
- Supportive extra reading
- Lecture powerpoints
- Recorded videos
- Analytical examples and worksheets
- Other: _____

7. Which of the following would help you prepare better for future online exams?

Check all that apply.

- Workshops with assessment coordinators
- Additional study skills sessions run by the department/college
- More on-line resources
- Better access to the University facilities
- Other: _____

8. Please could you make recommendations for how the University could improve preparing you for on-line exams?

9. In terms of your own preparation, which statement best describes your approach?

Mark only one oval.

- I was well prepared, I knew what and how to revise, I managed my time well
- I was somewhat prepared, I knew what and how to revise, but I could have managed my time better
- I didn't feel well prepared, I wasn't sure how to revise, I managed my time well, but just didn't know what to expect
- I didn't feel well prepared, I wasn't sure how to revise, I didn't really manage my time well

10. How did you prepare for your exams?

Check all that apply.

- Making notes and memorising facts, information and/or processes
- Made organised notes and learnt key concepts
- Wrote practice essays
- Discussed with peers
- Practiced past papers
- Worked through example problems/exercises/lab sheets
- Other: _____

11. How much time did you spend preparing?

Mark only one oval.

- Started a few weeks before the exam
- A few days of intensive preparation before the exam
- Last minute evening before organising notes
- Other: _____

12. Please can you give a short outline of how you prepared and revised

13. Did your preparation for online exams differ from the way you have prepared for exams in the past? If so please can you briefly describe this below

14. Did you experience any of the following difficulties during your revision and preparation time?

Check all that apply:

- Broadband issues
- Difficulty getting quiet study time
- Issues with resources (e.g. access to books, reading material)
- PC issues
- Illness
- Family/relationship issues
- No everything was fine
- Other: _____

15. Did you have any other concerns beyond the general exam anxiety? If so, please explain

The exam

This section aims to determine how you experienced the actual exam

16. Did you manage to complete the assessment in the recommended time?

Mark only one oval.

- Yes
 No

17. Did you feel your preparation and revision strategy effectively prepared you for your exams?

Mark only one oval.

- Yes
 No
 Maybe

18. If you answered no to the above, what would you change for the next session?

19. If you had a submission window (e.g., a 2 hour exam plus 1 hour extra to submit), did you use this to work on your assessment?

Mark only one oval.

- Yes, but only minor formatting changes
- Yes, I used all the given time
- No, the time was perfect and I managed to get everything done

20. Did you experience any of the following technical issues during the exam window?

Check all that apply.

- Internet/broadband failure
- Exam not uploaded
- Submissions issues
- Administrative errors in submission
- Extra time not provided for SpLD when required
- PC issues
- Errors in the exam script
- No I was fine
- Other: _____

21. Was your environment suitable for completing the exam?

Mark only one oval.

- Yes I have a quiet space with good internet access
- OK, I had some minor disturbances but overall it was fine
- I found it quite difficult to get a quiet space with good internet access, but managed in the end
- No, I struggled with constant disruption

22. Is there anything you want to add about your experience during the assessment?

Your overall learning experience

Here we want to know how these on-line exams have helped you to learn and progress and determine your overall experience.

23. On-line exams helped me to develop in-depth understanding of my topics

Mark only one oval.

- Strongly disagree
- Disagree
- Neutral
- Agree
- Strongly Agree

24. On-line exams helped me to develop rote learning and memorisation

Mark only one oval.

- Strongly disagree
- Disagree
- Neutral
- Agree
- Strongly agree

25. On-line exams helped me to develop long-term memory of my topic

Mark only one oval.

- Strongly disagree
- Disagree
- Neutral
- Agree
- Strongly agree

26. On-line exams helped me to identify relationships between different concepts

Mark only one oval.

- Strongly disagree
- Disagree
- Neutral
- Agree
- Strongly Agree

27. On-line exams have enabled me to evaluate and validate information

Mark only one oval.

- Strongly disagree
- Disagree
- Neutral
- Agree
- Strongly agree

28. Which of the following skills have on-line exams helped you to develop?

Check all that apply.

- Critical analysis
- Identifying and prioritising key information
- Answering questions coherently
- Organisation and time management
- Source and identify a range of relevant information
- Make effective use of library and online resources
- Correctly reference and paraphrase sources
- Problem solving
- Other: _____

29. Did your learning for on-line exams differ compared to traditional exams? If yes, please explain how.

30. How have on-line exams helped to prepare you for future study or employment?

31. Thinking about on-line exams, which of the following applies? I found them..

Mark only one oval.

- Less stressful than traditional exams
- Somewhat less stressful than traditional exams
- Just as stressful as traditional exams
- Somewhat more stressful than traditional exams
- More stressful than traditional exams

32. Is there anything you would like to add about how on-line exams have impacted your learning?

Thank you for completing this questionnaire. We are very grateful for your time and look forward to hearing all about your experience.

This content is neither created nor endorsed by Google.

Google Forms

Author details

Laura Roberts is a Professor in Biological Sciences who specialises in ecology and conservation. She is currently the Associate Dean for Education in the Faculty of Science and Engineering in Swansea University. Laura is a Principle Fellow of the Advanced HE and a Member of the Chartered Institute for Ecologists and Environmental Managers. Her pedagogic research interests are in student employability and assessment and feedback.

Joanne Berry is a Professor in Roman History. She specialises in the material culture of Roman Italy, particularly Pompeii and the other cities in the Bay of Naples. Jo is School Education Lead for Swansea University's School of Culture and Communication and the previous Dean of Assessment and Feedback. She is also a Principal Fellow of the Advanced HE and has been researching assessment and feedback in higher education, with a current focus on academic integrity.

Licence

©2023 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC-BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. See <http://creativecommons.org/licenses/by/4.0/>. Journal of Learning Development in Higher Education (JLDHE) is a peer-reviewed open access journal published by the Association for Learning Development in Higher Education (ALDinHE).