

Invigilated and Non-invigilated Online Assessments

Cost-effective, scalable online assessment solutions to assure academic integrity, privacy and equity of access: Towards a framework for success



Associate Professor Elaine Huber (Project Lead)

Academic Director, Business Co-Design, The University of Sydney Business School

Professor Lynne Harris

Head of Teaching and Learning, Chartered Accountants Australia New Zealand (CAANZ)

Professor Sue Wright

Head of Accounting Discipline Group, University of Technology (UTS) Business School, Sydney

Corina Raduescu

Senior Lecturer in Business Information Systems, The University of Sydney Business School

Dr Amanda White

Senior Lecturer and Deputy Head (Education), Accounting Discipline Group, Academic Integrity Officer, UTS Business School

Dr Andrew Cram

Lecturer in Educational Development, Business Co-Design, The University of Sydney Business School

Dr Sandris Zeivots

Lecturer in Educational Development, Business Co-Design, The University of Sydney Business School

Andrew Brodzeli

Senior Research Assistant, The University of Sydney Business School

Executive Summary

In the context of the rapid transition to fully online teaching and learning, the ABDC commissioned a project into online assessment in higher education in Australia. This report summarises the findings of that project, providing current evidence about the forms of online assessment in use in undergraduate and postgraduate business courses at Australian ABDC member institutions, and developing a framework to guide best-practice decision-making about online assessments.

We based our investigations on five design considerations for online assessments: the assessments must assure academic integrity, allow for the provision of quality feedback, support a positive learning experience for students, assure the integrity of student information and be delivered so that all enrolled students have an equal chance to complete the assessment successfully. Using input from a comprehensive literature review, the results of a survey of educators, and focus group discussions, we have extended these considerations to include authenticity. Our analysis confirmed that scale of delivery and resource limitations are broader and interrelated contextual factors that influence decisions about assessment design. These contextual factors also include institution policies and accreditation requirements.

This report contains summary information about the assessment types in use in business disciplines within Australia. We found that most survey respondents used written assessments and online exams / quizzes, with more than half also reporting that they used live or recorded presentations. The report also summarises the constraints and trade-offs identified in focus group discussions.

The outputs of this project include a summary diagram to assist future educators and accreditors in applying our framework of design criteria for online assessment solutions, and a description of our online portal that will enable our project to be current beyond 2022, as academics access our framework and share their best practice assessment designs. We conclude this report with our recommendations as to the way forward for future educators, researchers and peak funding bodies.

Acknowledgments

We would like to thank the Australian Business Deans Council for allocating funding to conduct this project. We also wish to thank our survey respondents and focus group attendees for their time to answer our questions and share examples of their practice.

This research was carried out with approval from the University of Sydney Human Ethics Research Committee (approval number: 2021/800).

TABLE OF CONTENTS

1. Introduction	5
1.1 Background and Aims	5
1.2 Our Approach.....	5
1.2.1 Definitions and Assumptions	6
1.3 Developing an Evaluation Framework for Online Assessment in Business Education	7
1.3.1 Literature Review.....	7
2. Data Collection: Evaluating the Framework	11
2.1 Survey.....	11
2.1.1 Survey Respondents	11
2.2 Online Assessment in Business — Survey Insights.....	12
2.2.1 What is Important for Choosing Online Assessments: Unprompted Open-ended Responses	12
2.2.2 What is Important for Choosing Online Assessments: Ratings.....	14
2.2.3 Forms of Online Assessment in Business	17
2.2.4 Invigilation of Online Exams and Quizzes	19
2.2.5 Relationship of Assessment Type to Accreditation Requirement, Assessment Weighting, Course Level and Size of Cohort	20
2.3 Applying Evaluative Framework to Online Assessment in Business	22
2.3.1 Survey insights: Ratings of Criteria	22
2.4 Focus Groups.....	24
2.4.1 Focus Group Participants.....	24
2.4.2 Focus Group Insights	25
3. Discussion and the Way Forward	29
3.1 Overview	29
3.2 Barriers to Innovation	30
3.3 Authenticity	31
3.4 Accreditation	31
3.5 Invigilation	31
3.6 The Way Forward.....	31
3.6.1 ABDC Cross-Institutional Leadership.....	31
3.6.2 Extension to Other Professional Disciplines.....	32
3.6.3 Extension to Student Perspective	32
3.6.4 More Research Funding for Good Online Assessment Design	32
3.6.5 Collaboration.....	32
3.7 Limitations.....	33
4. Project Outputs	34
4.1 Framework for Evaluating Quality Online Assessments.....	34
4.2 Online Portal	35
4.3 Good Practice Examples.....	36
5. References	38

List of Tables

Table 1. Professional body accrediting courses represented in the survey

Table 2. Examples of responses coded in each category

Table 3. Breakdown of accreditation requirement, assessment weighting, course level and size of cohort by assessment type (exam, written assignment, and other)

Table 4. Mean ratings of exams, written assignments, and other assessments

Table 5. Characteristics of Focus Groups

Table 6. Focus Group Findings

Table 7: Examples of participants assessments self-rated against our criteria.

List of Figures:

Figure 1. Frequency of criteria mentioned in open-ended responses when asked to list the criteria important when choosing an online assessment

Figure 2. Frequency word cloud of the 50 top words when asked to list the criteria important when choosing an online assessment

Figure 3a & 3b. Percent ratings of criteria used when deciding which online assessment(s) to adopt.

Figure 4. Forms of online assessment reported by survey participants

Figure 5: Forms of online assessment by accreditation status

Figure 6. Use of Individual and Group assessments

Figure 7. Invigilation approaches used for each type of online exam/quiz

Figure 8. Invigilation approaches used for exams/quizzes in accredited and non-accredited courses

Figure 9. Framework to evaluate quality online assessments.

Figure 10. Screen shot of the exemplar portal.

Figure 11. Screen shot of an example assessment in more detail

1. Introduction

1.1 Background and Aims

In 2020 Australia responded to the COVID-19 pandemic by imposing strict travel restrictions on international and domestic travel and by introducing social distancing in social environments, including higher education. Australian higher education providers had no choice but to rapidly transition from campus-based activities to online teaching and learning to ensure the students' participation and progress were not disrupted.

Prior to 2020, campus-based teaching and learning constituted most of higher education delivery in Australia. The main reasons for this include visa conditions for onshore international students that restricted online delivery to one-third of teaching in a course (National Code of Practice for Providers of Education and Training to Overseas Students, 2018), expectations of professional accrediting bodies regarding invigilated assessment, and concerns about academic integrity with online delivery (e.g., Holden et al., 2021). These factors were particularly important in business education where international students comprise a large proportion of enrolments and where many disciplines have professional accreditation (e.g., accounting)

The rapid, large-scale transition to online delivery in higher education that occurred in 2020 was quickly supported by research to assure the quality of online teaching, learning and assessment, and the national regulator has established a repository of resources to guide quality online delivery (Assessment integrity: Online learning good practice). The previous reliance by many providers on invigilated on-campus examinations has given way to increased usage of alternative modes of assessment, and research has demonstrated some pedagogical benefits of online assessment (Butler-Henderson & Crawford, 2020). However, this change is not consistent with the requirement by accreditation bodies such as CPA Australia and Chartered Accountants Australia and New Zealand for invigilated assessment in business education. While the exclusive use of online assessment that complied with COVID-19 public health measures but may not meet the expectations concerning invigilation was accepted by accreditors as a necessary interim solution, its longer-term acceptance in the post-pandemic period is not certain.

In this context, the aim of our project was two-fold: 1) to record and collate the types of online assessment in use in Australian business education; and 2) to develop a framework of key design considerations for educators to use for evaluating online assessments. The framework will support decisions about the overall quality of assessment types that may also contribute to the anticipated discourse about online assessment between educators, higher education providers and professional accreditation bodies.

Specifically, the framework was conceived to allow academics and other educational specialists to evaluate whether current or proposed online assessment practices are *scalable, cost effective*, and achieve five key *benefits*: (1) assuring academic integrity; (2) providing quality feedback; (3) enhancing student experience; (4) maintaining student information integrity; and (5) supporting equitable student access.

1.2 Our Approach

Our interdisciplinary team with expertise from academic and professional organisations took a research-driven approach to this project.

We built on our knowledge of business education and our experience to develop a range of benefits we believed were required for quality online assessments. We contextualised and refined these dimensions via a systematic literature review, and constructed our evaluation framework. The process of developing the evaluative framework is described in more detail in Section 1.3 below.

Next, we developed survey questions for business educators in Australia, based on our preliminary review of the literature. As soon as the survey was prepared, we applied for human ethics approval to collect data and disseminate findings (HREC number: 2021/800). We received the human ethics approval later than we had expected, on 9th November 2021, which was close to the end of semester and the end of year shutdown. Although the timing was unfortunate, we immediately deployed our online survey to business educators via business school deans.

The online survey was developed to explore the online assessments currently in use in business education in Australia. We collected 97 responses from academics and other educational specialists employed in ABDC affiliated institutions. We sought to understand the dimensions that these respondents considered most important in selecting assessment types, and to analyse how they ranked the dimensions included in our draft evaluation framework. Following the survey, we conducted four focus groups with a total of 19 participants. Some of the focus group participants indicated an interest in further discussion at the end of the survey and others were sourced from our networks. The focus groups enabled us to conduct more in-depth investigation of online assessment types against the five dimensions. We sought to refine and validate our framework by asking participants to apply and unpack its dimensions against specific assessment types. We wanted to clarify and fill any gaps in our assumptions.

The final step of our project is to disseminate the project findings and results through the development of an online portal. This is discussed further in section four.

1.2.1 Definitions and Assumptions

At the beginning of the study, we reviewed definitions of online assessment and agreed to use one definition to ensure clarity and consistency in our study. The outcome of our review was a modified version of Allan (2020)'s definition of online exam¹:

“online assessment’ refers to assessments (including non-graded hurdle tasks) which are mediated or facilitated by digital technologies and delivered online.”

We refer to *authentic assessment* throughout this report. Villarroel et al. (2018) describe three factors that must be present in authentic teaching and assessment scenarios.

“Realism involves linking knowledge with everyday life and work, *contextualisation* characterises situations where knowledge can be applied in an analytical and thoughtful way, and *problematization* invokes a sense that what is learned can be used to solve a problem or meet a need” (Villarroel et al., 2018, p. 841).

Throughout this study we have made the reasonable and implicit assumption that all assessments are designed in alignment with the intended learning outcomes of the course and programs in which they are employed. The principle of constructive alignment underpins this assumption:

¹ “high-stakes summative assessment events, mediated by digital technologies, often taking place in a defined place or time and under secure conditions (e.g. invigilation, restrictions on access to course materials, notes or communication).”

*“Assessment tasks and associated criteria must test student attainment of the intended learning outcomes effectively and at the appropriate level. Where learning outcomes state skills and attitudes as well as knowledge, this should be appropriately reflected in the chosen assessment methods. This is known as **constructive alignment**”* (Warwick Academic Development Centre, 2022).

1.3 Developing an Evaluation Framework for Online Assessment in Business Education

We first defined the design considerations and contextual factors that underpinned our proposed evaluation framework for online assessment in business education. We built on a 2019/2020 University of Sydney study of online exam systems that identified key benefits of online over invigilated paper-based exams (Bryant & Ruello, 2019).

Then we reviewed the education literature for examples of innovation in online assessment in practice that meet these design considerations and contextual factors. It became apparent that there are a limited number of such examples in business education, with examinations remaining the most frequently used assessment type. We utilised this literature review to develop a research-informed evaluation framework.

In the following section we present the results of our literature review regarding the:

- **five design considerations:** academic integrity, quality feedback, student experience, equity of access, and privacy / student information integrity
- **two contextual factors:** scalability and resourcing.

1.3.1 Literature Review

We explored a large selection of peer reviewed literature for papers and reports of relevance, identifying 67 articles published between 2011 to 2021 with a focus on online assessment. Below we provide a brief description of each design consideration and contextual factor, followed by a synthesis of the relevant literature. An annotated bibliography can also be found on our online portal.

Academic Integrity

This design consideration covers both security – whether the assessment assures against outsourcing, impersonation and assistance; and authenticity – both alignment and relation to practice and/or discourse in a discipline, profession or workplace.

Most of the online assessment literature focusses on academic integrity, understood as addressing academic dishonesty. These studies take different perspectives: system-wide studies of the education sector (Blankenberger & Williams, 2020), staff perceptions of cheating (Birks et al., 2020; Okada et al., 2019; Dawson, 2018; Atkinson 2016; Gehringer & Peddycord, 2013), student perspectives (Awdry 2021; Dendir & Maxwell, 2020) and those that include both staff and students (Reedy et al., 2021; Rolim & Isaias, 2019). At least one study focusses on the role of online assignment design (Harper et al., 2021), while others consider format protocols to address academic dishonesty (Butler-Henderson & Crawford, 2020; Munoz & Mackay, 2019; Bengtsson, 2019; Vos, 2015; Fask et al., 2014). Academic dishonesty may occur (a) due to assessment design flaws, use of generic questions, or a lack of surveillance or control of circumstances (although invigilation has a controversial role discussed further on under privacy), and/or (b) cognitive offloading where

students are dissatisfied with, or lack a supportive learning environment and where tasks have high weighting, or are mis-timed (i.e. too short or too long). Staff awareness of cheating differs by assessment type while students and staff both perceive academic dishonesty generally to be higher online.

Furthermore, authentic assessments have been shown to positively influence academic integrity by inherently reducing the likelihood of cheating through unique student responses (Dawson, 2020; Bretag et al., 2019). This is particularly relevant to business education where the literature shows a focus on using authentic assessments (e.g., James & Casidy, 2018; Vos, 2015; Ladyshewsky, 2015). Ellis et al. (2020) describe authentic assessment according to five factors: frequency, fidelity, complexity, impact and feed-forward. However, the authors challenge previous claims that authentic assessment design can assure academic integrity. For example, employing authentic assessment can cognitively overload students, thereby encouraging cheating behaviors.

Quality Feedback

This design consideration covers a number of considerations: capacity for provision of immediate feedback; feedback that encourages student-educator dialogue; feed-forward, defined as formative feedback towards later assessments; and enabling the provision of feedback through multiple formats not limited to written feedback.

Several studies discuss the provision of feedback within online assessment formats. Boitshwarelo et al. (2017) synthesise principles for effective delivery of online feedback in respect of online exams, recommending that online feedback about test performance be immediate and corrective. They recommend that academics should reference resources to reinforce learning and enhance learner self-regulation, and that feedback be used to improve future assessment design.

Other studies discuss the possibilities for new modes of feedback, i.e., audio, video and screencast technology enabled (Rolim & Isaias, 2019; Dawson & Henderson, 2017; Debusse & Lawley, 2016). Pitt and Winstone (2020) identify technology-enabled peer-to-peer and formative feedback, using learning analytics to generate individualised feedback and enhancing student motivation to engage with feedback. Others discuss the “dialogic feedback cycle” criteria for evaluating technologically delivered feedback through the provision of action points for feed-forward and encouraging student-educator dialogue throughout the process (Moscrop et al., 2017).

Debusse and Lawley (2016) report on their trial of automated feedback program SuperMarkIt for business students, finding students and educators agree on benefits such as quantity of feedback, legibility and personalisation. Other studies have provided a protocol that allows personalised feedback in practical authentic online assessment for pharmacy students, including a bespoke automated feedback platform, resulting in improved student experience and learning outcomes, and more consistent grading, while reducing overall teaching time committed (Ellis & Barber, 2016). Regarding multiple-choice question exams, the value of immediate and corrective feedback to enhance student engagement has been emphasized (Douglas et al., 2012).

Student Experience

We drilled down into the factors that create a positive experience for students during the completion of online assessment. For example, a positive experience may be how the assessment enables students to reduce cognitive load, stress and anxiety through good design and use of technology; how it enhances student motivation and concentration; or how it enhances convenience and comfort to students.

Several studies provide multifactorial discussions that interpret results of student surveys of their experience of online assessment, such as Cramp et al. (2019) who consider the value of reducing cognitive load, convenience, technical support, exam design and practice tools. Bin Mubayrik (2020) provides a literature review of studies with a 'student-centred approach' in online assessment, identifying six key aspects of student experience: motivation, classroom participation, task understanding, depth of learning, self-regulated learning and self-monitoring.

Other studies emphasise the connection between student satisfaction and higher-order cognitive engagement (Curtis et al., 2021; James & Casidy, 2018). Additionally, several authors discuss the causes of students' differential experiences apart from online assessment format, such as their motivation and career-orientation (Curtis et al., 2021; Rolim et al., 2019), and Bretag et al. (2019) identify the prevalence of contract cheating as an indication of student dissatisfaction with the broader teaching and learning environment.

A key determinant of improved student experience was ample communication of assessment task format and student expectations for those unfamiliar with new modes of online assessment (Bearman et al., 2020; Apps et al., 2020). Sullivan (2016) suggests that student experiences of online multiple choice tests are more positive when set formatively, with unlimited attempts allowed to facilitate greater and more self-regulated learning. Kolski (2018) emphasises pre-assessment familiarisation strategies such as non-weighted practice assessments.

Several studies consider negative aspects of student experiences, largely clustered around anxiety and technological issues associated with online assessment and consider measures to ameliorate these concerns (e.g., Moore, 2018; James, 2016). Linden and Gonzalez (2021) recommend enabling students to interact with technical support via a chat window during the assessment. Butler-Henderson and Crawford (2020) provide recommendations for assessment portal interface design, such as incorporating an autosave feature. Myyry and Joutsenvirta (2015) promote open-book/open-web formats to enhance student feelings of self-efficacy and reduce assessment anxiety.

Equity of Access

This design consideration examines a variety of important factors including ability to offer live technical support; to enable different assessment conditions to meet individual student needs; to offer flexible access (for geographically dispersed students); to avoid potential discrimination during automated grading/feedback; and ensuring that access to hardware and software is fair, particularly if information regarding technical requirements was not available before decisions about enrolment were made.

Several organisations produced valuable reports about equity of access in online assessment space given their focus on equity in higher and tertiary education (Austin et al., 2021; ADCET, 2020; Stone, 2017). Despite these reports and focus, equity of access is an under-researched area. Studies that focus on student perspectives share common access concerns regarding the technological and home environment suitability for completing online exams, and the fairness of differential access to technology and comfortable home conditions (Kharbat & Abu Daabes, 2021; Rahman, 2021; Linden & Gonzalez, 2021).

ADCET (2020), Atherton (2021) and Austin et al. (2021) provide insight into how students with disabilities are uniquely affected by the change to online assessment in response to COVID. These include higher education providers' failure to consider access requirements, failure to adequately equip staff to cater for students with disabilities, and student experiences of isolation and disproportionate financial pressures. Tai et al. (2022) considers four dimensions of differential equity

in tertiary education: students with disability, students from regional, rural and remote areas, students with low socioeconomic status, and students who were first-in-family to obtain tertiary education. The authors found these students perceived positively the following mechanisms that could increase their access requirements: increase staff to student support, simplify the special consideration process, provide flexible timing of assessment, and schedule open-book and/or practical and authentic assessment.

Several studies concerned with online assessment design provide recommendations for measures to improve accessibility (e.g., Lloyd et al., 2021; Stone, 2017; Ladyshevsky, 2014). For example, Stone (2017) recommends early contact of educators with students regarding online assessments, presenting information in multiple ways, or using early formative assessment to build academic expectations.

Privacy / Student Information Integrity

This design consideration covers the need to reduce the likelihood of collection or breaches of personal details such as demographic and biometric data. It also covers the reduction or avoidance of sharing of student generated content. We also acknowledge here the tension that exists between academic integrity and privacy/security of student information. Often technology, such as artificial intelligence used to assure online exam invigilation, requires provision of personal data, identifying information that may be susceptible to security breaches. The literature also reveals tensions between student experience and online proctoring, some claiming both positive and negative perceptions (Jaap et al., 2021; Milone et al., 2017).

Student surveys have revealed a common student concern regarding privacy and especially in relation to online proctoring (Kharbat & Abu Daabes, 2021; Linden & Gonzalez, 2021; James, 2016). Okada et al. (2019) raise the issue of biometric data leaks because of the production and storage of such data. Coghlan et al. (2021) provide a philosophical discussion of the ethics of online exam supervision technologies including in respect of privacy and discuss the inadvertent capture of personal or sensitive information, and the sharing of such data with third parties and subsequent leakage.

Contextual Factors: Scalability and Resourcing

These factors were initially envisaged as overarching considerations that moderate the application and implementation of the abovementioned five dimensions during assessment design.

Scalability

Online marking platforms have successfully enabled scaled-up feedback and grading mechanisms (Bearman et al., 2020). However, they also open up the possibility of failures in automated marking tools (Dawson & Sutherland-Smith, 2018; Steel et al., 2019). Carless et al. (2017) identify the potential of online technology for enabling scaled-up assessment-for-learning (i.e., employing educators' understandings of student progress to inform their teaching). Finding ways to improve student learning experiences while reducing teacher time commitment and resourcing, such as through using multi-modal feedback, the employment of e-portfolios and curriculum mapping are some of these potentials.

Resourcing

This factor refers to staff time and financial resources on assessment design. For example, when designing assessments, consideration of the time and cost of assessment handling, marking, and

grade processing are important concerns. Depending on factors such as the size of the student cohort, different decisions will be made about assessment design. Adequate resourcing is the primary constraint to achieving robust academic integrity and assessment authenticity in scaled-up online programs (Apps et al., 2020; Birks et al., 2020; Ellis et al., 2020; Cramp et al., 2019). For example, Holden et al. (2021) suggest that the relatively higher cost of live online invigilation discourages its uptake in favour of automated invigilation. Other studies reveal that lack of institutional resourcing of adequate staff training can constrain the efficacy of online assessment (Okada et al., 2019; Rolim & Isaias, 2019).

2. Data Collection: Evaluating the Framework

We collected data through an online survey and four focus groups from participants with experience designing and delivering online assessments.

2.1 Survey

We gathered data on the following:

- Demographic information
- Online assessment types used in business education
- Reasons for designing online assessments (free form)
- Exemplars / well-designed online assessments (free form)
- Ranking of the framework dimensions using the exemplars

2.1.1 Survey Respondents

A total of 97 respondents completed the survey, of whom 91 were affiliated with a university. Respondents were recruited from business faculties around Australia and represent 15 different disciplines. Accounting is the discipline with the strongest representation, followed by management, finance, human resources, marketing and economics.

The respondents were asked to identify all of the roles they held in their institutions. The majority of respondents were unit coordinators (n=85; 88%) and/or lecturers (n= 69; 71%). Other categories were program coordinators (n=36; 37%), tutors (n=32; 33%) and education designers (n=15; 15%). There was one associate dean and one head of school in the sample.

Regarding course accreditation by a professional membership body, around two-thirds of the respondents (n=66) indicated that the courses they worked on were accredited, and 18 respondents indicated that the courses they contributed to were not accredited.

Respondents were asked to indicate which professional organisation accredited the courses they worked on, where courses may be accredited with more than one professional organisation. This information is summarised in Table 1. Not all respondents understood their course accreditation status fully, with 15 respondents unsure about accreditation status, and a total of 20 respondents either unsure about accreditation status or what accreditation bodies were involved if it was accredited. Further analysis showed that almost all (n= 17) of these unsure respondents had roles as Subject and/or Program Coordinators.

Table 1. Professional body accrediting courses represented in the survey

Professional Accrediting Body	Frequency
Certified Professional Accountants (CPA) Australia	38
Chartered Accountants ANZ (CAANZ)	37
Australian Human Resources Institute	9
Financial Advisor Standards and Ethics Authority (FESEA)	9
Australian Computer Society (ACS)	5
Chartered Financial Analyst (CFA)	5
Law Societies (State based)	5
Other*	12
Unsure	5

* * Association to Advance Collegiate Schools of Business (AACSB); Australian Marketing Institute (AMI); Australian Property Institute; Actuarial Sciences (State based); Public Relations Institute of Australia; Engineers Australia; and Public Relations Institute of Australia

2.2 Online Assessment in Business — Survey Insights

We invited participants to provide their opinions of online assessment before asking them directly about what we had included in our framework, so that we were not prompting them. In the section below, we present the respondents' (unprompted) rankings of what is important to them when considering online assessment.

2.2.1 What is Important for Choosing Online Assessments: Unprompted Open-ended Responses

We asked participants to 'List the general criteria² that are important to you when choosing which online assessment to adopt'. Responses were coded independently by two members of the research team and the coding was verified by two different members. Examples of each category are shown in Table 2.

Table 2. Examples of responses coded in each category

Category	Examples
Academic integrity	"To act as an assessment where it meets student identification verification. Completed individually to meet the above requirement. Timed assessment in a short window of time to restrict academic misconduct issues."
Authenticity	"This makes it important to frame questions so that they don't rely on access to information but rather focus on understanding and interpretation."

² Initially we used the term "criteria" in the survey and the focus groups to cover what we now call key design considerations and contextual factors. In this report, we have faithfully recorded the terminology used with our respondents and participants, while using our updated and more accurate terms as much as possible. In some sections we refer to dimensions as a generic term for key design considerations and contextual factors.

Category	Examples
Student experience	"Must be manageable for students without the technical problems we encountered with AI invigilation."
Addressing subject & learning outcomes	"That the assessment type (e.g. quiz, report, discussion) be aligned with/ appropriate to the learning outcome/s being assessed."
Assessment is varied, randomised	"There is variety and choice of assessment tasks offered to students." "Multiple choices questions are able to be shuffled."
Staff experience/ marking	"How much workload is involved in designing, deploying and marking the different types of assessment?"
Equity	"Supporting access to education for students in remote locations or other situations preventing access to classroom."
Assessment is time restricted	"Ensuring time limits are enforced."
Institution	"The criteria I adopt is the one set by the faculty."
Technology	"By far the most important criterion is whether the online assessment format/technology that I have can accommodate the kinds of questions that are most appropriate in my subject."
Accreditation	"Accrediting body policy - ignorance for appropriate/alternate assessment design."
Student Information Integrity	"I am also deeply uncomfortable with online invigilation, and so design all of my final assessments so that students are not recorded or observed through their computers by the university while they take my exam."
Quality Feedback	"Students are provided with action focused feedback (only for formative assessment, no feedback is provided for summative assessment)."

Figure 1 shows the frequency of different items coded by the research team in open-ended responses to the request 'List the general criteria that are important to you when choosing which online assessment to adopt'.

Of the 97 respondents 69 (71%) answered this question and respondents were able to list as many items as they wished. Consequently, the figures above represent the number of participants mentioning each item in their responses.

When asked to provide their most important considerations when adopting an online assessment, respondents were most commonly concerned about **academic integrity** (n=32; 46% of responses) and **authenticity** (n=23; 33% of responses), **student experience** (n=21; 30%) and then **addressing subject & learning outcomes** (n=19, 28%).

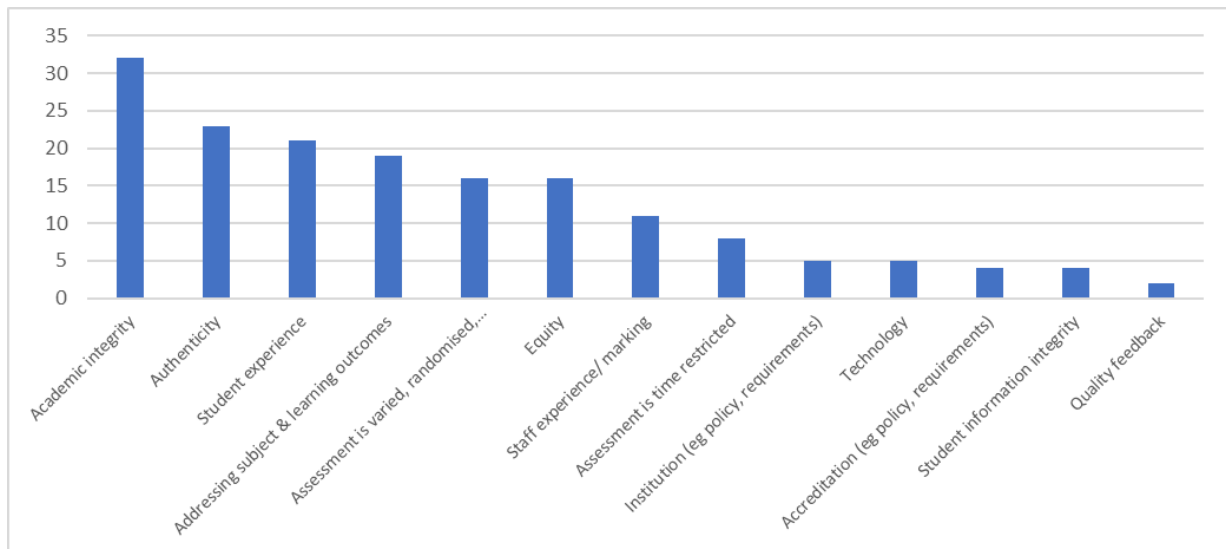


Figure 1. Frequency of criteria mentioned in open-ended responses when asked to list the criteria important when choosing an online assessment

Of the remaining design dimensions, **student information integrity** (n=4; 5% of responses) was mentioned next and **quality feedback** (n=2; 3% of responses) least most frequently identified. We note with the grouping of **equity** (n=16, 23%) that it included not only equity of access but also other equity issues and accessibility.

Of the other influential contextual factors, **institution** (e.g., policy, other requirements) (n=8; 12% of responses), **technology** (n=5; 7% of responses), and **accreditation** (n=4; 6% of responses) were identified by some respondents.

Meeting or addressing the subject demands and learning outcomes (n=19; 28% of responses) was also a commonly mentioned consideration. This was not included in our evaluative framework because our implicit assumption is that this is an essential element of all assessment design decisions. This was not, however, made explicit in the survey. It is therefore not surprising that constructive alignment between assessments and learning outcomes was mentioned as an important consideration by many respondents.

A frequency word cloud of the open-ended responses is also provided in Figure 2.

2.2.2 What is Important for Choosing Online Assessments: Ratings

We then asked survey respondents to rate the importance of our suggested design considerations (academic integrity, quality feedback, student experience, information integrity and equity of access) and contextual factors (scalability and cost effectiveness, plus others such as aligning with institutional assessment culture) as well as the key underlying principle of assuring learning outcomes. Participants rated a series of statements using a Likert scale ranging from 'not important at all' (1) to 'very important' (5).

In Figures 3a and 3b, we present the responses for each design consideration, showing the proportion of respondents that rated each item by the five levels of importance. We have grouped responses in coloured boxes, where green includes our original five design considerations, orange includes additional considerations and red indicates intended learning outcomes.

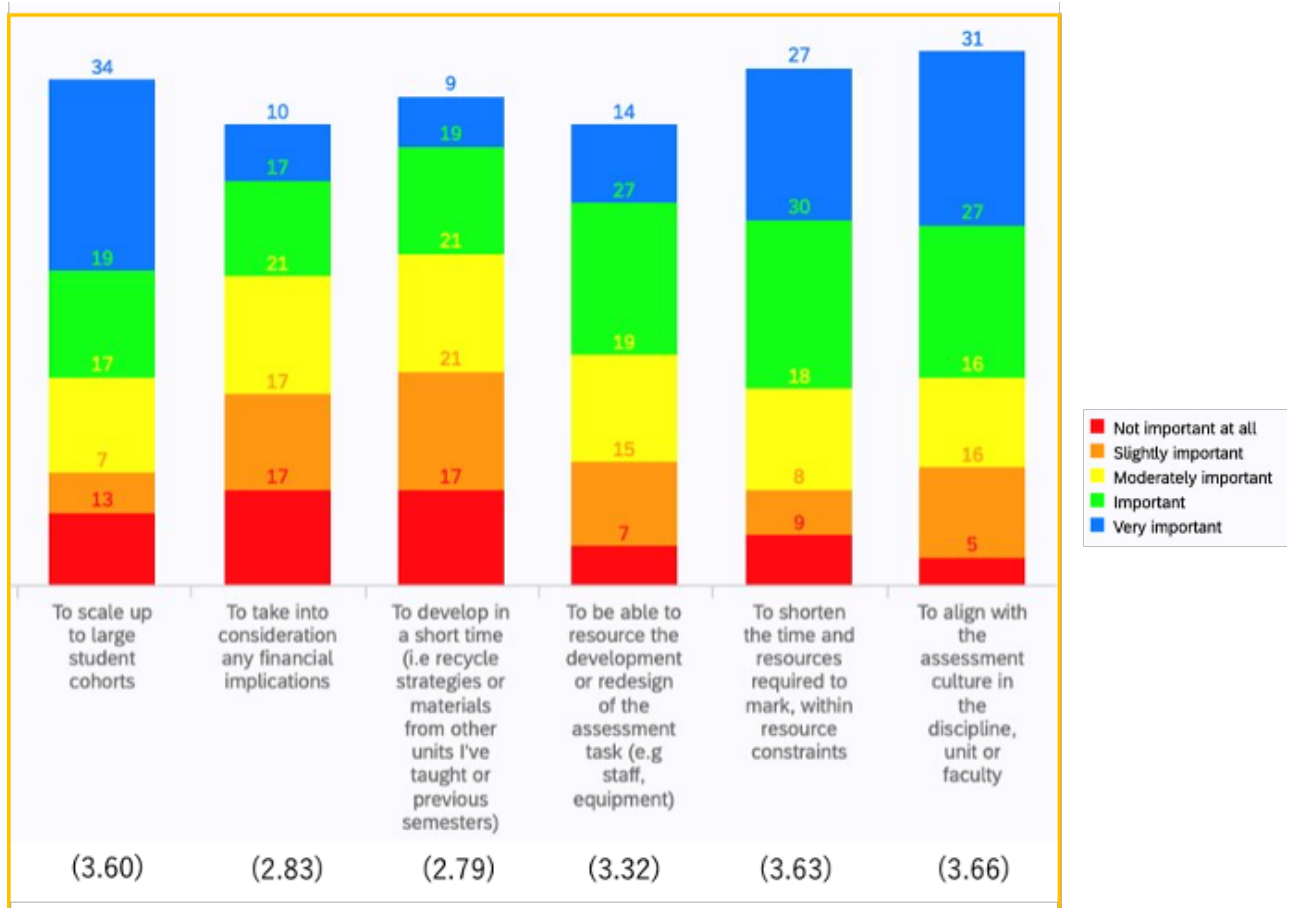


Figure 3b. Percent ratings of criteria used when deciding which online assessment(s) to adopt

The results show that each of the key design considerations included in our initial evaluation framework (academic integrity, quality feedback, student experience, student information integrity, equity of access) were ranked highly, and all had mean ratings above 4.00 (“Important”). Taken together with the open-ended responses, these survey responses provide a preliminary validation of the criteria we included in our framework. Not surprisingly, ‘mastering intended learning outcomes’, which is fundamental to assessment design decisions, was also rated as important by participants (M= 4.44). Additional design considerations concerning scale of delivery, resourcing and institutional policy considerations were all ranked below 4.00 (Mean ratings ranged from 2.79 to 3.66; see Figure 3b).

Also noteworthy was the variation between the responses to the prompted question (Figure 3) versus the unprompted question (Figure 1). Some criteria were rated as important when prompted but were not mentioned in the unprompted responses. These included quality feedback, student information integrity and equity of access. Authenticity was commonly ranked as important when unprompted, although it was not an explicitly prompted option in the survey. This led us to consider including authenticity as an additional key design consideration in our framework. We discuss this further in section 3.

2.2.3 Forms of Online Assessment in Business

This section provides a snapshot of the forms of online assessment being used by survey respondents. Participants were asked to indicate as many forms of assessment as they had used from a list provided (n=97) (see Figure 4).

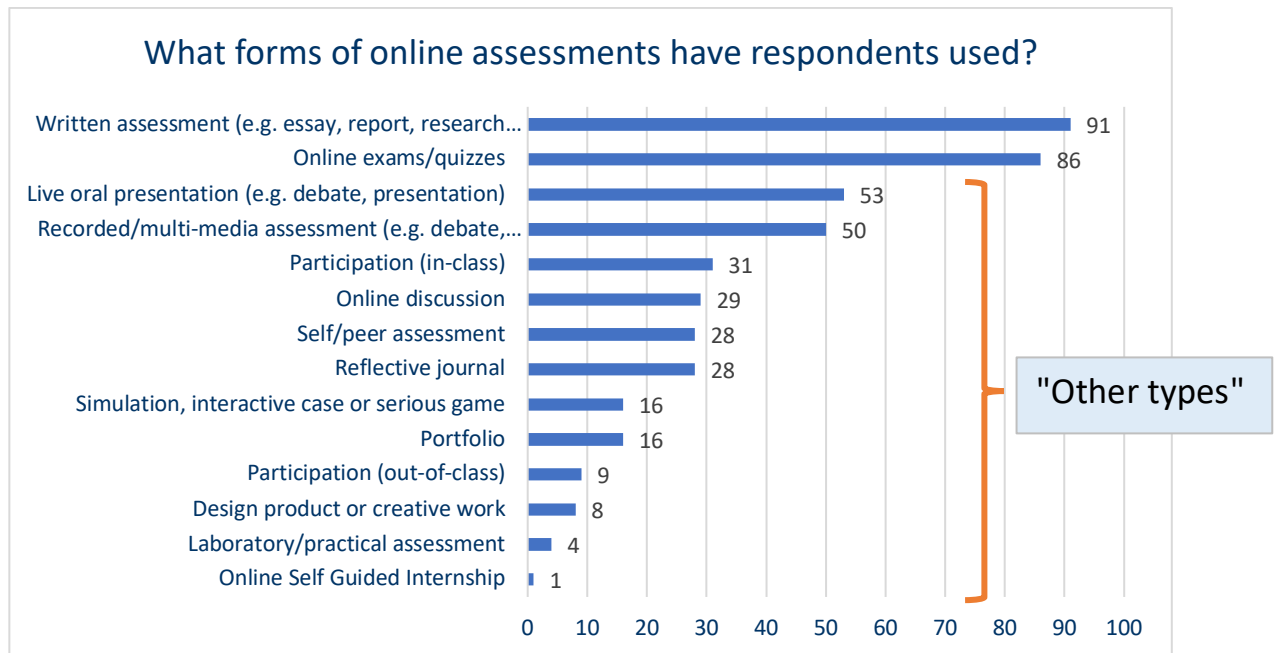


Figure 4. Forms of online assessment reported by survey participants

Figure 4 indicates that written assessments and online exams or quizzes were reported as being used by almost all respondents, while slightly over half of the respondents had used live oral presentations and recorded/multi-media assessments.

Around a third of respondents reported using in-class participation assessment, self/peer assessment, online discussion and reflective journals. Fewer respondents reported using portfolios, simulations, out-of-class participation, design or creative works, or laboratory/practical assessments. The initial list of assessment types developed for the survey was found to be very comprehensive, with only one assessment type added to our list from the responses: 'Online self-guided internship'.

For the purpose of further analysis, we categorized the assessments as written assessments, online exams/quizzes and "other", where "other" comprised a range of assessment types that were less commonly used (see Figure 4).

The use of each form of assessment (exams/quizzes, written, or other) was generally similar across accredited courses and non-accredited courses (see Figure 5).

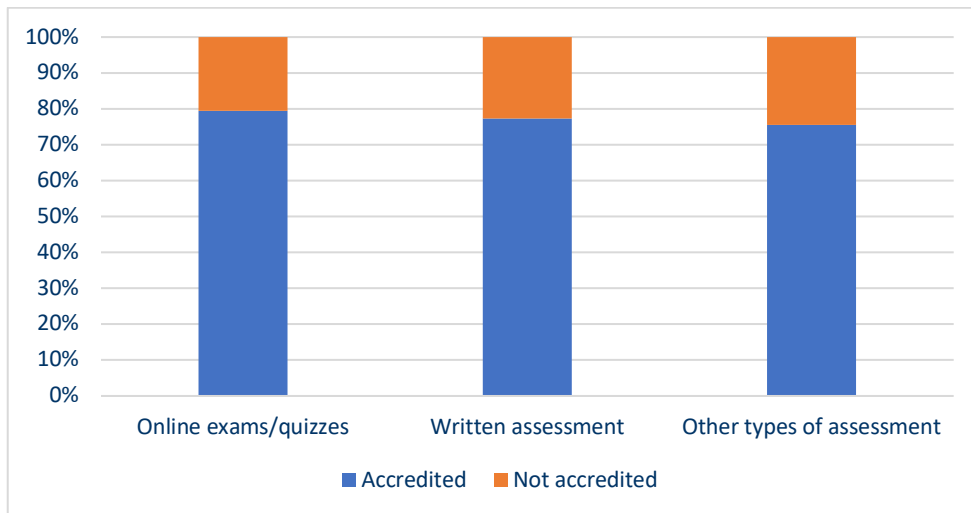


Figure 5: Forms of online assessment by accreditation status

Online exams and quizzes were overwhelmingly used as individual assessments, and only six respondents (7%) indicated they had used exams/quizzes as group assessments (see Figure 6). In contrast, 52 respondents (57%) who had used written assignments had set these as group assessments. Where respondents reported using ‘other’ forms of assessment, 79% of the assessments were set for individuals, while 51% of assessments were set for groups. Note that some respondents had set an assessment type for both groups and individuals so the options are not mutually exclusive.

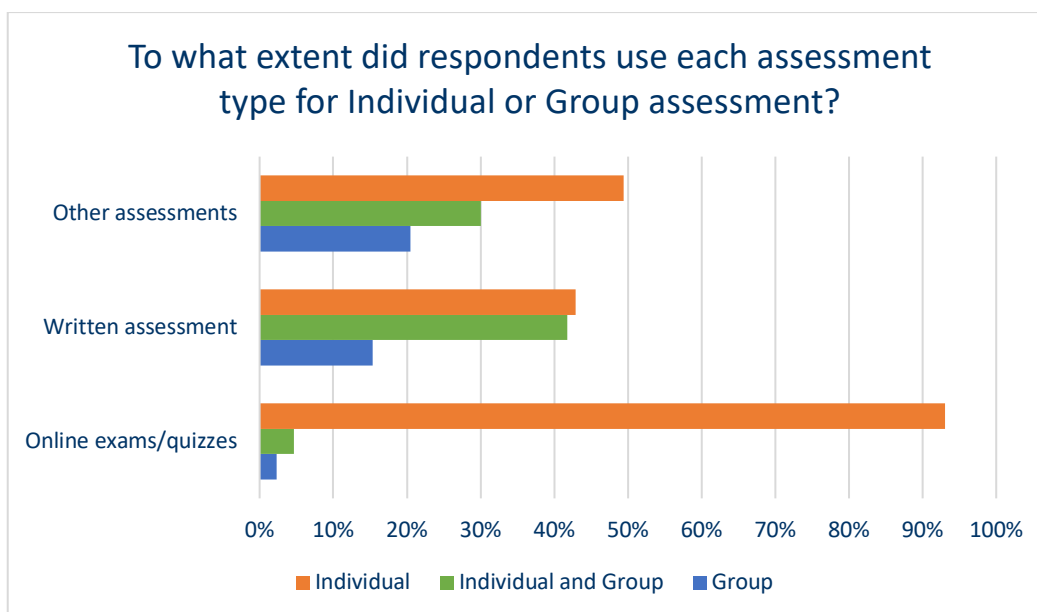


Figure 6. Use of Individual and Group assessments

2.2.4 Invigilation of Online Exams and Quizzes

The 85 respondents who indicated they had used online exams/quizzes were also asked about invigilation and what invigilation methods they used for different question types. Respondents could select multiple options in this question as they may have run several different online exams/quizzes using different invigilation approaches.

Figure 7 shows the relative use of each invigilation method for each type of online exam/quiz question. Within this sample, the largest observable difference was that exams with extended written answers were more likely to be set as 'take home' exams and less likely to be 'computer-invigilated' compared to the other types of exams. On the other hand, the invigilation approaches for exams involving numerical calculations and multiple-choice questions followed a similar pattern, and there was a similar prevalence of people-only invigilation across each type of online exam/quiz.

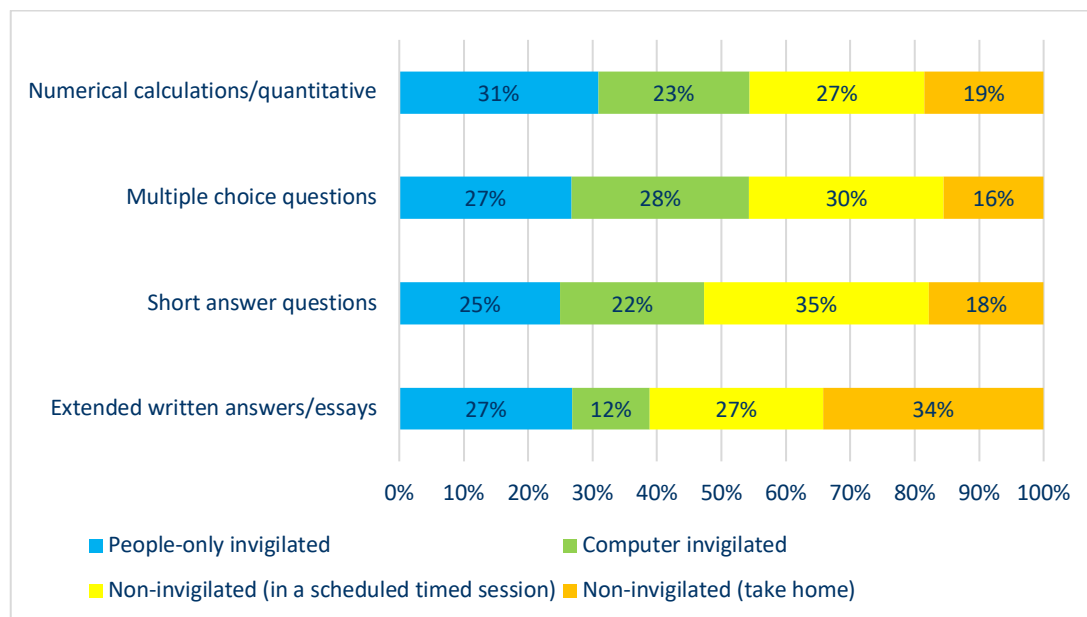


Figure 7. Invigilation approaches used for each type of online exam/quiz

Analysis was also performed on whether the accreditation status of a course influenced the selection of invigilation type (Figure 8). The results suggested that accredited and non-accredited courses were both equally likely to use invigilation (either computer or people-only) for their exams/quizzes. However, in this sample a difference was observed in approaches for non-invigilated exams/quizzes: accredited courses were more likely to use scheduled timed sessions, and non-accredited courses were more likely to use take home exams.

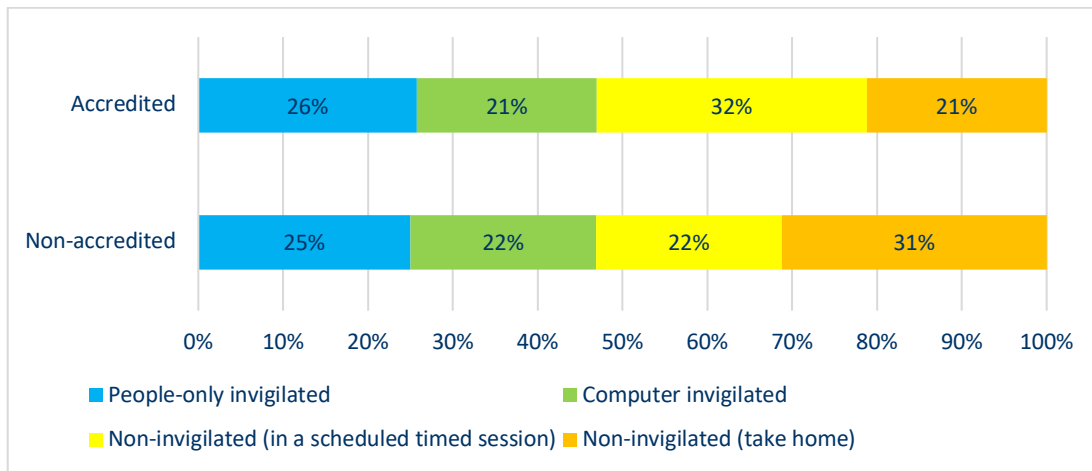


Figure 8. Invigilation approaches used for exams/quizzes in accredited and non-accredited courses

The respondents who had used computer invigilation methods (n=35) were asked to provide further details on what methods they had used. The most prevalent computer invigilation method was artificial intelligence invigilation (17), followed by teacher-facilitated zoom invigilation (7), lockdown browsers (4), live invigilation (3), a combination of lockdown browser and live invigilation (1), and other non-specified options (3).

2.2.5 Relationship of Assessment Type to Accreditation Requirement, Assessment Weighting, Course Level and Size of Cohort

We also looked at the relationship between assessment type and whether the assessment was required for accreditation, the assessments weighting, whether the assessment was part of an undergraduate or a postgraduate course and the size of the cohort where the assessment was being used (see Table 3).

Table 3. Breakdown of accreditation requirement, assessment weighting, course level and size of cohort by assessment type (exam, written assignment, and other)

Statement	Exam (n=36)	Written (n=31)	Other (n=30)	p-value	Effect size
Accreditation				0.281 ^{ns}	0.177
Assessment is required for accreditation	18 (47.2%)	9 (29%)	10 (33.3%)		
Assessment not required for accreditation	14 (38.9%)	16 (51.6%)	14 (46.7%)		
Weighting				<0.001	0.455
Assessment is weighted ≥ 31% (major)	24 (66.7%)	20 (66.7%)	5 (17.2%)		
Assessment is weighted ≤ 30% (minor)	12 (33.3%)	10 (13.3%)	24 (82.8%)		
Course level				0.052 ^{ns}	0.245
Undergraduate	29 (78.4%)	20 (64.5%)	15 (50%)		
Postgraduate	8 (21.6%)	11 (35.5%)	15 (50%)		
Cohort size				0.003 ^{**}	0.319
Small (1-29)	4 (10.8%)	7 (22.6%)	4 (13.3%)		
Medium (30-99)	13 (35.1%)	5 (16.1%)	19 (63.3%)		
Large (100-249)	7 (18.9%)	12 (38.7%)	3 (10%)		
Very large (250+)	13 (35.1%)	7 (22.6%)	4 (13.3%)		

* p<.05; **p<.005; ns: non-significant

The chi-square test of independence was used to examine the relationship between the pairs of variables in Table 3.

Accreditation Requirement

There was no significant relationship between whether an assessment was required for accreditation and assessment type.

Weighting of Assessment

The relationship between weighting and assessment type was significant where 'other' assessments were more likely to be lower weighted and exams and written assignments were more likely to be higher weighted.

Course Level

Course level was not significantly associated with assessment type.

Cohort Size

Cohort size was significantly associated with assessment type where 'other' assessment types were more likely to be used in small and medium cohorts (see Table 3).

2.3 Applying Evaluative Framework to Online Assessment in Business

2.3.1 Survey insights: Ratings of Criteria

Participants were asked to select one well-designed online (summative) assessment that they had used and to rate that assessment according to its capacity to satisfy a series of criteria presented as statements. The ratings were on a scale from 'Not at all' (1), 'Very little' (2), 'Somewhat' (3) to 'To a great extent' (4).

For the purpose of analysis, and consistent with the classification in Section 2.2.3 above, the assessment types were classified as written assessment (n=31), online exams/quizzes (n=36) and other (n=30). Analysis of variance (ANOVA) was used to examine differences between the mean ratings of each statement for the three groups (see Table 4).

Table 4. Mean ratings of exams, written assignments, and other assessments

Statement	Exam (n=36)	Written (n=31)	Other (n=30)	p-value	Effect size
Academic integrity					
Ensures academic security	3.19	2.97	3.60 ^{Rnk}	0.002**	0.379
Authenticity					
Aligns with tasks commonly done in a discipline, profession or workplace	3.36	3.71	3.47 ^{Rnk}	0.017*	0.283
Is performed in similar conditions to the tasks conducted in a discipline, profession or workplace	2.78	3.45	3.07	0.008*	0.329
Involves a degree of complexity that requires an inquiry-based approach	2.94	3.29	3.57	0.014*	0.256
Involves application beyond the educational setting	2.81	3.39	3.67	0.002**	0.384
Involves scaffolded self-assessment	2.44	3.71	2.77	0.614 ^{ns}	0.106
Student experience					
Enhances convenience and comfort for students	3.56	3.14	3.41	0.158 ^{ns}	0.215
Enables students to reduce cognitive load	3.29	2.66	2.74	0.017*	0.302
Enhances student motivation and concentration	3.41	3.52	3.70	0.184 ^{ns}	0.175
Reduces student stress and anxiety during the assessment	3.20	2.75	3.00	0.143 ^{ns}	0.211
Reduces likelihood of technical complications	2.85	3.34	3.04	0.152 ^{ns}	0.214
Student information integrity					
Reduces the likelihood of collection or breaches of personal student details, such as demographic and biometric data	3.39	3.15	3.43	0.480 ^{ns}	0.140
Reduces or avoids the sharing of student generated content	3.14	3.03	3.50	0.113 ^{ns}	0.203
Equity of access					
Enables different assessment conditions to meet individual student needs	3.61	3.38	3.21	0.099 ^{ns}	0.228
Offers live technical support	3.09	2.50	2.62	0.061 ^{ns}	0.264
Enables flexible access to assessment (e.g., geographically dispersed students)	3.83	3.74	3.79	0.777 ^{ns}	0.072
Enables the provision of immediate feedback	2.88	2.73	2.83	0.899 ^{ns}	0.049

Statement	Exam (n=36)	Written (n=31)	Other (n=30)	p-value	Effect size
Enables the provision of feedback through multiple formats	2.34	2.82	2.86	0.194 ^{ns}	0.208
Encourages student-educator dialogue	2.59	2.93	3.41	0.001**	0.396
Facilitates formative feedback toward later assessments	3.04	3.19	3.54 ^{Rnk}	0.092 ^{ns}	0.249
Enables academic peer feedback on the assessment	2.34	2.72	2.86	0.276 ^{ns}	0.185
Responds to student perceptions of the assessment	2.84	3.00	3.27	0.188 ^{ns}	0.192
Scalability					
Provides mechanisms for feedback at scale (e.g. feedback comment banks, automated grading, or AV feedback)	3.24	2.56	2.40	0.003**	0.353
Enables the collection of education data to inform further assessments	3.44	2.66	2.79	0.003**	0.340
Expedites managing assessment	3.61	2.43	2.45	<0.0001	0.690
Expedites grading	3.60	2.17	2.60	<0.0001	0.638
Resourcing					
Increases financial cost	3.18	2.89	3.10	0.579 ^{ns}	0.120
Increases time and resources to develop the assessment	2.08	2.20	2.64	0.089 ^{ns}	0.237
Increases time and resources required to implement and administer the assessment	2.56	2.33	2.54	0.701 ^{ns}	0.091
Increases the marking time and resources	2.92	2.31	2.70	0.101 ^{ns}	0.238
Influential factors					
Aligns with institutional policy	3.23	3.06	2.62	0.078 ^{ns}	0.244
Is required for accreditation	2.71	2.26	1.92	0.040*	0.274
Continuity of assessment	3.34	2.77	2.40	0.003**	0.365

^{Rnk} Ranked ANOVA; * p<.05; **p<.005; ns: non-significant

Academic Integrity

There was a significant difference between the assessment type groups on academic security, where the means show that ‘other’ assessment types were rated highest.

Authenticity

There were significant differences between the groups in ratings for four of the five statements (see Table 4). The pattern of means indicates that ‘other’ assessments are rated highest for: ensuring academic security; degree of complexity that requires an inquiry-based approach; and involving applications beyond an educational setting. Written assessments were rated highest for: aligning with tasks commonly done in a profession or workplace; and for being performed in similar conditions to the tasks conducted in a discipline, profession or workplace.

Student Experience

The only significant difference between the groups for these five statements concerned enabling students to reduce their cognitive load, where the pattern of means shows that online exams were rated highest.

Student Information Integrity

There were no significant differences between groups for ratings of either of these two statements.

Equity of Access

There were no significant differences between groups for ratings of any of the four statements in this category.

Quality Feedback

For the six statements in this category, the only significant difference between groups concerned encouraging student-educator dialogue where the pattern of means shows that online exams were rated highest.

Scalability

There were significant differences between the groups in ratings of all four statements in this category, where the pattern of means indicates that online exams are rated highest on each of the statements.

Resourcing

There are no significant differences between groups for ratings of any of the four statements concerning resourcing.

Influential Factors

There were significant differences between groups for ratings of Continuity of assessment, which refers to maintaining a consistent assessment format and/or structure across time for a particular subject or unit and for continuity. For both of these, online exams were rated highest, followed by written assessments, and finally other assessment types.

2.4 Focus Groups

We conducted four focus groups, in which we asked the participants to comment on the application of the five design considerations to one or more different assessment types as described in Figure 4.

2.4.1 Focus Group Participants

Each focus group had four to six participants (total of 19 participants). Each focus group was facilitated by two members of the project team. The participants were 10 women and 9 men from four states/territories and from a range of universities. Table 5 summarises the characteristics of focus group participants.

Table 5. Characteristics of Focus Groups

	Focus Group 1	Focus Group 2	Focus Group 3	Focus Group 4
Date	13/12/2021	25/01/2022	01/02/2022	03/02/2022
Focus: application of the criteria to	Online assessment in general	Individual reflective journals	Group debate and peer feedback	Industry workshop and live oral assessment
Number of participants	4	6	5	4
Women:Men	1:3	3:3	4:1	2:2
Type of institutions represented*	Go8: 3 ATN: 1	Go8: 3 RUN: 1 Other: 2	Go8: 4 ATN: 1	Go8: 3 ATN: 1
State/Territory	ACT: 1 NSW: 1 Vic: 2	NSW: 5 Vic: 1	ACT: 1 NSW: 3 SA: 1	NSW: 2 SA: 1 Vic: 1

* ATN: Australian Technology Network; Go8: Group of Eight; RUN: Regional Universities Network; Other: Not affiliated with Go8, RUN or ATN

2.4.2 Focus Group Insights

We present our findings from the focus groups in table format (Table 6), highlighting the key tradeoffs between our five design considerations and additional contextual factors (scalability, authenticity, and academics' individual concerns).

Table 6. Focus Group Findings

Criteria	Tradeoff
Academic Integrity	<p>TRADEOFF BETWEEN INTEGRITY USING INVIGILATION AND AUTHENTICITY / STUDENT EXPERIENCE</p> <ul style="list-style-type: none"> It is difficult to assure integrity in the online environment without some form of invigilation and identity verification. Most participants related to assessments where students present material synchronously (e.g. exam or live presentation). <p>TRADEOFF BETWEEN INTEGRITY WITHOUT INVIGILATION AND SCALABILITY</p> <p>Where invigilation is not possible or is not used, the most common solutions to assure academic integrity involve:</p> <ul style="list-style-type: none"> authentic assessment that applies skills and knowledge to case-based examples; and/or staff assessment of the process of developing the assessment as well as the product (e.g., observations of collaboration on a group project within class); and/or assessment of unique / within class experiences (e.g. individual student reflection or reflection with a classmate on an in-class activity; peer review of the process of group work).

Criteria	Tradeoff
	<p>TRADEOFF BETWEEN ASYNCHRONOUS ASSESSMENTS AND INTEGRITY / EQUITY / FEEDBACK</p> <ul style="list-style-type: none"> Assessments that are not conducted at the same time for all students (eg within tutorial activities across a week or across the term) create particular issues for academic integrity due to 'leakage'. They also therefore pose challenges for equity and provision of feedback, as those undertaking assessments earlier and later have differential access to feedback. These issues are exacerbated in the online environment compared to campus-based delivery due to ease of information sharing.
Student experience	<ul style="list-style-type: none"> We note that participants did not readily think of the student experience. Student experience needs to be considered at the course level as well as the unit level. While it is good to teach students how to undertake a reflective assessment task, this only needs to be done once in the course, in one subject. This category is impacted by INSTITUTIONAL CONTEXT (e.g. if students cannot be required to turn cameras on, this can result in low engagement or response from students which affects their experience). <p>TRADEOFF BETWEEN STUDENT EXPERIENCE AND ACADEMIC INTEGRITY</p> <ul style="list-style-type: none"> Short timeframes around assessments can leave students feeling they have no option but to engage in cheating. Hybrid classes can detract from student experience (but may be needed for opportunities to identify students and thus reduce Academic Integrity issues). <p>TRADEOFF BETWEEN STUDENT EXPERIENCE AND EQUITY OF ACCESS / CAREER CONCERNS</p> <ul style="list-style-type: none"> Regardless of mode of delivery, if assessment experiences are not consistent then they are perceived as unfair and may lead to complaints which may impact on academics' 'reputation'. <p>TRADEOFF BETWEEN STUDENT EXPERIENCE AND PRIVACY</p> <ul style="list-style-type: none"> Students may not have good internet access if they are internationally or remotely located, or cannot afford it, or are living in share accommodation (Privacy). <p>TRADEOFF BETWEEN STUDENT EXPERIENCE AND SCALABILITY</p> <ul style="list-style-type: none"> Cost and scalability were raised in relation to assessment decisions. Academics are interested in quality student experience, but scalability can affect it "drastically".
Quality Feedback	<p>Quality feedback is:</p> <ul style="list-style-type: none"> actionable consistent timely for the recipient (individual and broader group)

Criteria	Tradeoff
	<ul style="list-style-type: none"> defensible. <p>TRADEOFF BETWEEN QUALITY FEEDBACK AND SCALABILITY</p> <ul style="list-style-type: none"> If academics aren't trained to provide defensible feedback this can lead to complaints. The solution may be rubrics, but they must include all important aspects in the criteria. Should there be feedback on final exams? Some institutions' practice is to not give feedback due to resource requirements. Others have introduced final exam feedback. Innovative assessments like student presentations require a lot of time to provide quality feedback. <p>TRADEOFF BETWEEN QUALITY FEEDBACK AND PRIVACY</p> <ul style="list-style-type: none"> The use of oral/audio/video feedback rather than written feedback from academics and practitioners may provide efficiencies but requires staff training for consistency/appropriate tone etc. <p>TRADEOFF BETWEEN QUALITY FEEDBACK AND EQUITY OF ACCESS / STUDENT EXPERIENCE</p> <ul style="list-style-type: none"> Peer feedback can be useful but it must be clear that there is a purpose to using peer feedback and students must be trained to provide peer feedback. If purpose is unclear and/or students are not prepared and/or peer feedback is not anonymous, the feedback is unlikely to be valid. <p>FOR ACADEMICS, TRADEOFF BETWEEN QUALITY FEEDBACK AND CAREER PROGRESS</p> <ul style="list-style-type: none"> Giving good quality and honest feedback affects how tutors are perceived (and impacts on their student ratings).
Equity of access	<p>TRADEOFF BETWEEN EQUITY OF ACCESS AND STUDENT EXPERIENCE</p> <ul style="list-style-type: none"> Internet access particularly impacts synchronous activities with video/audio interaction and/or use of web-based tools Timezones / locations (particular problem for group work, often solved by (a) scheduling classes to conduct groupwork and/or (b) specifying tools to be used for meetings). Presentations require skills that vary by culture, gender, personality etc. This is an accepted aspect of assessments assuming that the assessment is selected to assure key intended learning outcomes. As accessibility adjustments may remove option of live presentations, there needs to be consideration of how these are assured for all students. It is important to use tools that are easily and universally accessible. Equity requires that students have opportunities to practice with the technology required for assessments (e.g. invigilation software). <p>TRADEOFF BETWEEN EQUITY OF ACCESS AND SCALABILITY</p> <ul style="list-style-type: none"> One solution to equity of access issues is to provide options/choice in assessments. This may be appropriate for small cohorts only.

Criteria	Tradeoff
Privacy	<ul style="list-style-type: none"> • Privacy must be addressed/assured at an institutional level. • It must be clear for personal material (e.g., identifying information, video recordings) how long / in what form / where / how security is assured. • Reflection exercises create specific problems for privacy. An issue was reported around responding to safety concerns raised by reflections — there is a need to make sure this is notified ahead of the assessment.
Authenticity	<ul style="list-style-type: none"> • Involving industry people is important (authenticity, networking) but if they carry out assessment they must be trained, including key policy frameworks (e.g., academic integrity, equity and diversity).
Scalability and Cost	<p>TRADEOFF BETWEEN SCALABILITY AND COST EFFECTIVENESS</p> <ul style="list-style-type: none"> • Online marking is much more efficient, reducing time in accessing exams for marking (logistics). • The relative resourcing of large classes is smaller and therefore introducing anything new/ innovative is less likely if resources are scarce.
Academics' individual Concerns	<p>This group of findings speaks to the participants' concerns that assessment changes can impact not only on their promotion or career prospects but even more simply on job security.</p> <ul style="list-style-type: none"> • Issues and / or changes in large classes may be more 'visible' (both to students and institutions). It can be perceived as more risky to the academic to innovate. • Regardless of mode of delivery, if assessment experiences are not consistent then they are perceived as unfair and cause complaints (Equity of Access). • Giving good quality and honest feedback affects how tutors are perceived (and student ratings). • There is a lack of institutional recognition of the emotional labour involved in online teaching and providing personal contact with students to maintain engagement and student satisfaction – this is also an increased workload. • There is a lack of students' recognition of a reasonable response time – both need to be reframed and expectations made clear. • Institutions are seen as focusing on student support during the pandemic (not surprising) but at the expense of staff, as academics are also subject to many of the same pressures as students (e.g. working from home; children in remote learning). • Some solutions to providing quality assessment involved selecting tutoring staff with fewer family responsibilities (e.g. males with no children).

3. Discussion and the Way Forward

3.1 Overview

Since 2020 there has been a rapid transition to fully online forms of assessment in higher education within Australia in response to social distancing requirements prompted by the COVID-19 pandemic. This project provides current evidence about the forms of online assessment in use in undergraduate and postgraduate business courses in Australian ABDC member institutions. It also examined the key design considerations applied by educators when selecting assessment types for online delivery, which has been incorporated into existing research findings to develop and refine a framework to guide best-practice decision-making about online assessments.

Starting with a literature review, we identified five considerations for online assessments, at the same time recognising that how these considerations are operationalised will be influenced by the broader context in which the assessment is situated. The five design considerations are that the assessments must assure academic integrity, allow for the provision of quality feedback, support a positive learning experience for students, assure the integrity of student information and be delivered so that all enrolled students have an equal chance to complete the assessment successfully. We identified scale of delivery and resource limitations as broader and interrelated contextual factors that influence decisions about assessment design to meet the five key considerations mentioned above. We noted that scale of delivery varies markedly between units of study depending on factors such as whether they are foundational undergraduate units that form part of generic undergraduate degrees, more bespoke, later year undergraduate electives, or components of specialised postgraduate courses. We also noted that resource availability, while partly determined by scale, is also related to characteristics of individual institutions.

The literature review also informed the survey that we developed to identify the assessment types in use in business disciplines within Australia. We found that the majority of survey respondents used written assessments and online exams / quizzes, with more than half also reporting that they used live or recorded presentations. In addition, a range of other forms of online assessment were also in use (see Figure 4). We confirmed that academics who responded to our survey were using the five key design considerations and were also using authenticity as an additional key design consideration.

From the focus group discussions, we discerned some of the processes that respondents went through as they applied the key design considerations to choose and/or develop assessments. It was clear that the broader context of scale and resource availability constrained design decisions as did assessment weighting and the assessment policies of the institution, where small variations in policy direction had major impacts on the capacity of assessment to validly assure the key design considerations. For example, a policy variation in whether students could be required to have cameras on during online classes had a major impact on the academic integrity of assessments that incorporated class-based participation or presentation. As well, participants explained 'trading-off' elements of the key design considerations against each other in developing assessments and confirmed the impact of the broader contextual factors on the choice of assessments that led to prioritising of certain design elements in particular contexts. For example, low weighted, formative assessments were more likely to prioritise student experience and deprioritise academic integrity compared to more heavily weighted, summative assessments where academic integrity was prioritised.

3.2 Barriers to Innovation

It was clear from our survey findings and focus group discussions that innovation in assessment was not widespread among Australian ABDC member institutions where traditional forms of assessment such as written reports, exams and quizzes continued to be commonly used. This may be because these forms of assessment are tried and tested, work well and translate readily to the online environment. Alternatively, there may have been little appetite for making major changes to assessment over the past two years in the context of the many challenges associated with COVID-19 including the short timeframes for moving assessment to fully online delivery, unpredictable staff and student absences due to illness, and institutional budget constraints and staffing changes due to forecasts of much reduced international enrolments. From our focus groups, it is clear that academics viewed assessment innovation as a high stakes / high risk undertaking, with potential for impact on job security and career progression. Focus group participants also highlighted the impact of the rapid move to conducting teaching and assessment online on academic staff workloads particularly for those who had not previously taught online where training in new systems was required. For focus group participants, innovation was often seen as a trade-off of student experience and student pastoral care which were particularly prioritised in 2020. As fully online delivery becomes the new normal in a post-pandemic higher education environment and the conditions mitigating against innovation are declining, we argue that the time for supporting innovative online assessment design has arrived. We present our framework, portal, and good practice exemplars to assist academic staff decisions about assessment.

3.2.1 Broader Contextual Factors

As noted above, our framework includes scale of delivery and resource limitations as broader and interrelated contextual factors that influence decisions about assessment design. For large cohorts, many decisions about assessment design were taken to manage resource limitations that can be measured in financial terms, staff availability, turnaround time and a range of other metrics. In particular, we noticed that assessment decisions for large cohorts tended to be linked to feedback and became a trade-off decision for individual, nuanced feedback within these resource constraints. In terms of solutions, there is support in the literature for a range of technology-enabled individual and group feedback methods using audio and video in addition to text (Dawson & Henderson, 2017; Ellis & Barber, 2016; Pitt & Winstone, 2020) which can provide more individualised feedback at scale. For low weighted assessments, peer assessment was used by our focus group participants, although the value of this appeared to depend on the preparation that students were given for peer assessment activities and whether the activities were integrated into the learning outcomes of the unit of study.

Our focus group participants came from a range of institutions, including institutions affiliated with the Australian Technology Network, the Group of Eight, the Regional University Network and unaffiliated institutions. The policies of the institutions necessarily impact the way in which assessments are designed and implemented. Examples of this include variations in whether students are required to turn cameras on in scheduled classes, whether invigilated exam software is available or mandated, and workload models which impact on resourcing. We note that these variations make it difficult to provide sector-wide guidance. However, our framework is a first step in this direction, bearing in mind that institutional policy directions are understood as broader contextual factors impacting assessment decisions.

3.3 Authenticity

The assessment types that we collectively designated as ‘other’ had in common that they assessed performance directly, compared to written assessments, exams and quizzes that often reflect descriptions of how an activity should be performed. Performance-based assessments can be both authentic, in that they can require the demonstration of skills such as communication appropriate to a specified audience, and can mitigate academic integrity risks as the individual presenting the assessed work can be clearly identified. Our focus group discussions cautioned that the value of performance-based assessments in the online environment varied according to institutional policies and scale as they relied on individuals being identifiable to assessors. There is a plethora of literature on authentic assessment and some studies have found that authenticity alone cannot assure academic integrity (Bretag et al, 2019; Elias et al., 2020). In particular, many students need to have the ‘authenticity’ scaffolded and supported which in turn requires resourcing.

3.4 Accreditation

Many participants were not aware of whether the courses they contribute to are accredited or not. This may be because accreditation requirements are not a major concern when designing assessments, or because the departmental, school or faculty guidelines for assessment are aligned with accreditation requirements so that these requirements are not ‘visible’ to academic staff selecting and designing assessments.

3.5 Invigilation

There were some differences observed in our data on the selection of invigilation methods according to the type of question being asked and the accreditation status of the course. However, it was not clear whether there was any leadership being enacted in this regard and how decision making was occurring. Participants indicated that decisions about invigilation could be difficult due to a perceived trade-off between academic integrity concerns and concerns over privacy of student data.

It appears to be a widely held urban myth that (invigilated) exams are required for accreditation. Furthermore, it has been found that some people have preconceived ideas around exams and invigilation by people (Bryant & Ruello, 2019). Given the vast uncertainty over the past two years, it is likely that academics are sticking to what they know, simply transposing the exam online and continuing to invigilate.

Institutional policies need to make assessment design requirements clear to enable teachers to make consistent judgements and allow more creative and innovative assessment design. We know that accrediting bodies require verification. We hope this study will provide options to academics to consider ‘performance’ types of assessment as alternatives to invigilated online exams.

3.6 The Way Forward

3.6.1 ABDC Cross-Institutional Leadership

Our findings show that broad contextual factors, including scale, resource limitations and institutional policy settings, necessarily impact the decisions that academics make about assessment design. However, these factors are seldom a focus of research in the educational literature which provides direction about best practice to support student learning and are outside the control of individual academics designing assessments. Modifying these broader factors requires institutional

and disciplinary leadership, which the ABDC can provide cross-institutionally. This may assist individual faculties liaising with their institutions in situations where institutional policies impact assessment design decisions.

3.6.2 Extension to Other Professional Disciplines

This project focused on the work being undertaken in business faculties and the sample was relatively small. Given the wealth of research and quality innovation in design and delivery of teaching and assessment being undertaken in other professional areas with accreditation requirements (e.g. medical education - Jaap et al., 2021; Pharmacy – Milone et al, 2017; Psychology – James, 2016), further work that extends the survey to other professional disciplinary areas is recommended.

3.6.3 Extension to Student Perspective

Given the short timeframe for this project, it was not possible to include student perspectives. This is an important limitation, and we argue that including student perspectives would greatly enhance the utility of the framework and should be explored in future studies.

3.6.4 More Research Funding for Good Online Assessment Design

This is the beginning of an investigation onto good online assessment design and not the end. Some evidence-based practice is beginning to emerge in the literature from the past two years of assessing during a pandemic, particularly considering the perceived need for online invigilation. This project adds to this knowledge pool. We strongly recommend funds are invested into more awards and grants for work in this area, to encourage more innovation and wider dissemination of good practice.

For example, it is timely to reconsider and interrogate the long-standing use of summative assessment practices such as exams particularly in light of the large scale move to online learning and teaching during the pandemic. More investigation is needed into 'other' innovative assessments.

A consideration that emerged unexpectedly in our focus groups concerned the stress experienced by academic staff in relation to online assessment and workload. Focus group participants raised a range of issues about the ways in which requirements of online assessment may be inconsistent with meeting the needs of learners. These concerns were exacerbated by the work environment where job security was not assured. Such concerns cannot be alleviated with advice about assessment design and require separate investigation.

3.6.5 Collaboration

Throughout our focus group discussions, we observed that most assessment design decisions were being made in isolation. Of course, changes are reviewed by curriculum committees and discussions take place with directors regarding program alignment, but the choices, details and implementation happen individually on a unit or subject level.

Our focus groups provided a model for cross-institutional discussions of assessment design, and we strongly recommend continuing facilitated discussions amongst peers, both within and across

institutions. Such conversations can support innovation and help lower the perceived risk involved in individual coordinators making assessment design decisions.

3.7 Limitations

Most importantly, the absence of the student perspective from an investigation of assessment design, where students are fundamentally involved in the assessment process, limits our findings and further work in this area should prioritise including the student voice.

Due to the initial delay in obtaining ethics approval, our survey commencement was delayed. This meant that data collection occurred at the end of the academic year when many academics were focusing on their teaching priorities or were taking annual leave. This is one reason for a low number of participants in our survey. Another reason could be the indirect method of contacting participants. The ABDC emailed our survey to Business Deans, but from our networks we soon realised that many colleagues had not seen the survey. Finally, the impact of survey fatigue cannot be overlooked, given the proliferation of studies conducted on online teaching and learning in recent years.

4. Project Outputs

4.1 Framework for Evaluating Quality Online Assessments

Having tested our initial design considerations with the participants of this study, we have developed a visual presentation of our framework to encourage its application by educators, and potentially, accreditation bodies. As discussed previously in this report, we have added authenticity alongside our original five design considerations. We observe that two of our proposed criteria (scale and resourcing) are predetermined by the context in which the design and selection of online assessment types occurs. As a result, the framework consists of six key criteria and four contextual factors that are integral to evaluating and creating quality online assessment designs.

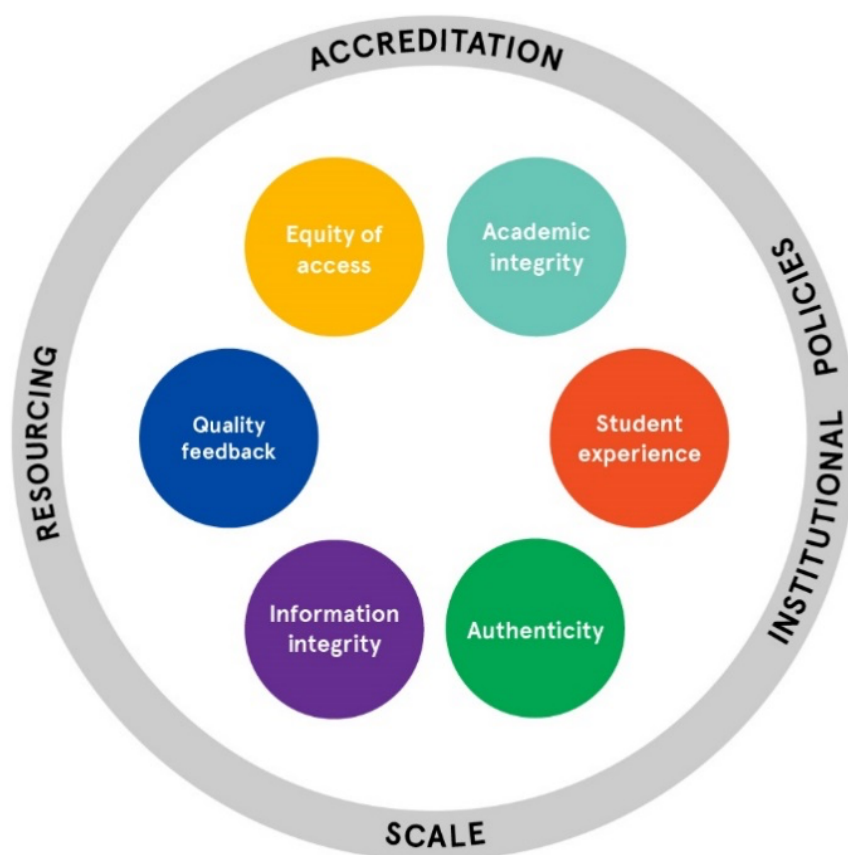


Figure 9. Framework to evaluate quality online assessments

We envisage this framework can be used in multiple ways, including the following:

It can be used to evaluate existing assessments, for an individual unit by the coordinator, or for a course or program as part of a general review of assessment design. It provides assurance over the design considerations that have been included in the evaluation. See Table 7 for some examples.

It can be used to document assessment practice and trade-offs between design considerations. This may open up conversations about the tradeoffs inherent in assessment design, and the pressures that exist in certain contexts.

It can be used to design new online assessments or redesign existing ones. A proposed change to assessment can use the framework to guide and demonstrate the reasons for or impact of the change.

4.2 Online Portal

An online portal has been created to showcase the outputs of this project including the report itself, our framework for evaluation of online assessments, and exemplars of online assessments and how they rate against the evaluation framework. The portal will also link to presentations of the report and publications that come from the findings.

<https://bizonlineassessment.com/>

The objective of the portal is to provide educators with inspiration to alter assessments from more traditional tasks such as essays and online exams to more innovative forms of assessments based on exemplars (some examples of these are provided in this report in Section 4.3 below and the remainder in our online portal). It provides a one stop shop for exemplars of assessments suitable to Business schools and faculties. The exemplars of assessment in the online portal can be searched by any specific characteristic of the evaluation framework (such as high on academic integrity, student equity, or privacy) and also by type of assessment (such as online exam, debate, group project, presentation). The portal also has the capacity to invite educators to submit other assessments to be included in exemplars for visitors to access.

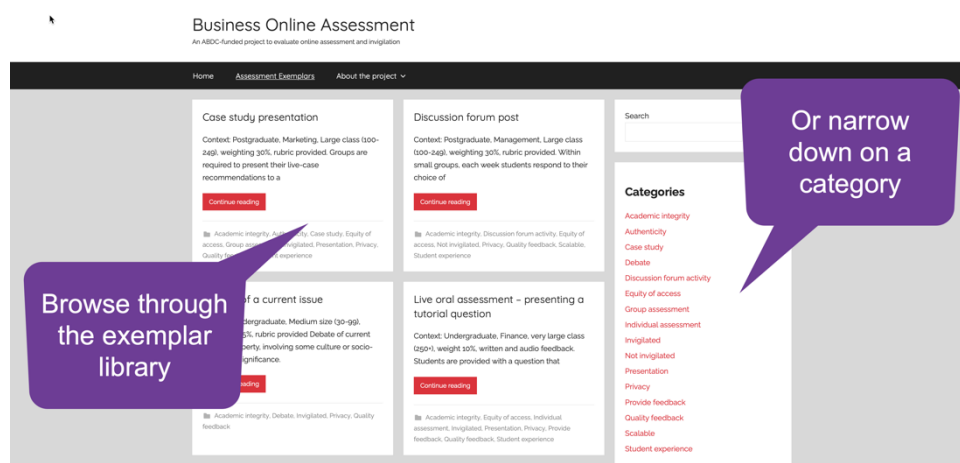


Figure 10. Screen shot of the exemplar portal

Within each exemplar is information provided by survey participants about online assessments being used throughout Australia higher education institutions.

Live oral assessment – presenting a tutorial question

Context: Undergraduate, Finance, very large class (250+), weight 10%, written and audio feedback.

Students are provided with a question that they must present to the tutorial class in the following week. Students are required to use slides and provide a detailed explanation. The question is relevant to the tutorial in which they have been allocated.

Students complete this assessment once during the semester.

Students must also, in different week, provide feedback on another student's presentation. Students are made aware that they must act as a reviewer one week prior to class. The student must write up their comments and send these with a suggested assessment mark for the presenting student to their tutorial via email.

Rated highly for	Rated moderately for	Did not rate strongly for
<ul style="list-style-type: none"> * academic integrity * student experience * privacy * providing quality feedback 	<ul style="list-style-type: none"> * equity of access * scalability 	<ul style="list-style-type: none"> * cost to implement

■ Academic integrity, Equity of access, Individual assessment, Invigilated, Presentation, Privacy, Provide feedback, Quality feedback, Student experience

Descriptions of exemplars

Ratings of exemplars against our framework

- Academic integrity
- Authenticity
- Case study
- Debate
- Provide feedback
- Quality feedback
- Scalable
- Student experience

Figure 11. Screen shot of an example assessment in more detail

4.3 Good Practice Examples

In our survey, we asked participants to share their examples of assessments they felt were well designed. We present these examples on our online portal in the form of a database that will be searchable by users. This includes participants' self-assessment against our framework to enable users to search for examples meeting particular criteria that align with their context. For completeness we also include some of the issues participants told us that they faced with implementing these assessments. Table 7 shows a sample of these examples. The remainder can be found on the portal.

Key:

The percentage given to each exemplar's rating with respect to the evaluative criteria were determined by, firstly, attributing values 0-3 according to how much the respondent agreed with the statement (0 = not at all, 1 = a little, 2 = moderately, 3 = to a great extent), secondly, aggregating the response scores for each statement within the evaluative criterion, and thirdly, excluding from the total any statements to which the response was 'not applicable or unsure'. The total maximum score for academic integrity would therefore be 18, as there are six statements that make up that criterion, and each statement has a maximum possible score of 3. These scores were then converted to a percentage in order to demonstrate equivalence between criterion scores where statements may have been excluded due to inapplicability to the particular assessment exemplar. For example, If one statement within the academic integrity criterion was considered inapplicable to a particular exemplar assessment, while the remaining statements were all agreed with 'to a great extent', then the score would be 15 (instead of 18), and the corresponding percentage would be the same (i.e. 100%). The only exception to this method was the 'cost' criterion, for which the attributed values were inverted, (0 = to a great extent, 1 = moderately, 2 = a little, 3 = not at all) to maintain the positive relationship between scores and performance. This was because the statements constituting the cost criterion were negatively framed, rather than positively framed, as for the other criteria.

Maximum possible scores for each criterion: academic integrity (18), student experience (15), privacy (6), equity of access (12), quality feedback (18), scalability (12), cost (12). The total scores would of course be lower if the respondent chose 'N/A' or 'unsure' for any of the options. The percentages however are representative.

Table 7: Three examples of participants' assessments self-rated against our criteria that were utilised in the focus groups

Assessment descriptor	Category / type	Context of use	Academic integrity	Student experience	Privacy	Equity of access	Quality feedback	Scalability	Cost	Challenges
Reflective e-portfolio journal Consolidation of the semester long experience, exploring the transformation from being newcomers to the subject (i.e., marketing) to where the students are at the end of the subject.	Reflective journal Individual Final	Undergraduate Marketing Medium cohort (30-99) 40% Weighting Rubric feedback Not required for accreditation	17/18 94%	8/15 53%	3/6 50%	7/9 78%	6/12 50%	6/9 67%	9/12 75%	The main challenge was the lack of integration between the portfolio host software and the learning management system (LMS), particularly the inability to mark them on LMS as the submissions being locked by the portfolio software. This creates unnecessary communication, confusing students regarding their submissions, and additional back and forth work for academics.
Debate of current issue of culture and socio-economic significance in groups of three students. It involves peer feedback as well as an assessment of each team member.	Debate (Group), Peer feedback	Undergraduate Property Medium cohort (30-99), Weighted 25%, Rubric and audio feedback, Required for accreditation	14/18 79%	4/15 27%	5/6 83%	7/12 58%	16/18 89%	3/12 25%	7/12 58%	One challenge is requiring students to debate issues that might include culture or social-economic sensitivity, as it will inevitably place stress on students. This type of assessment is appropriate for a capstone subject.
Industry hosted workshop and live oral assessment – professional skills development (e.g., building rapport, teamwork and professionalism) in groups. Students work in class, then simulate a client/practitioner meeting using a case study, the industry practitioners take on the role of client. At the end students submit a group report with case study solutions and a reflection on the task. Both academic and industry feedback is provided.	Live oral assessment Group	Undergraduate Accounting Medium cohort (30-99) 30% weighted Rubric and free comment and industry feedback Not required for accreditation	16/18 89%	10/15 67%	6/6 100%	8/12 67%	15/18 83%	11/12 92%	5/12 42%	Having a simulated meeting of a long duration is challenging due to time pressures. Students struggle to display all their soft skills in the time frame. Students also face difficulty in their groups due to dynamics and different working styles. Technical difficulties are also a challenge, although this could be due to lack of practice beforehand.

5. References

- ADCET. (n.d.). *Beyond COVID-19: The impact of COVID-19 on disability practitioners in the tertiary education sector: Survey report | VOCEDplus, the international tertiary education and research database* (p. 35). University of Tasmania. Retrieved March 16, 2022, from <https://www.voced.edu.au/content/ngv:87721>
- Allan, S. (2020). Migration and transformation: A sociomaterial analysis of practitioners' experiences with online exams. *Research in Learning Technology*, 28(0), Article 0. <https://doi.org/10.25304/rlt.v28.2279>
- Apps, T., Beckman, K., & Bennett, S. (2020). Challenges with Designing Online Assessment to Support Student Task Understanding. In M. Bearman, P. Dawson, R. Ajjawi, J. Tai, & D. Boud (Eds.), *Re-imagining University Assessment in a Digital World* (Vol. 7, pp. 247–262). Springer International Publishing. https://doi.org/10.1007/978-3-030-41956-1_17
- Atkinson, D., Nau, S. Z., & Symons, C. (2016). *Ten Years in the Academic Integrity Trenches: Experiences and Issues*. 27, 12.
- Austin, K., Ford, E., & Glencross, S. (2021). *Challenges to access and equity in the Australian higher education context* (Perspectives on the Challenges to Access and Equity in Higher Education across the World in the Context of COVID., p. 23). National Education Opportunities Network. <https://research.usc.edu.au/esploro/outputs/report/Challenges-to-access-and-equity-in/99570208502621>
- Awdry, R. (2021). Assignment outsourcing: Moving beyond contract cheating. *Assessment & Evaluation in Higher Education*, 46(2), 220–235. <https://doi.org/10.1080/02602938.2020.1765311>
- Bearman, M., Dawson, P., Ajjawi, R., Tai, J., & Boud, D. (Eds.). (2020). *Re-imagining University Assessment in a Digital World* (Vol. 7). Springer International Publishing. <https://doi.org/10.1007/978-3-030-41956-1>
- Bengtsson, L. (2019). Take-Home Exams in Higher Education: A Systematic Review. *Education Sciences*, 9(4), 267. <https://doi.org/10.3390/educsci9040267>
- Bin Mubayrik, H. (2020). New Trends in Formative-Summative Evaluations for Adult Education. *SAGE Open*, 10(3), 215824402094100. <https://doi.org/10.1177/2158244020941006>
- Birks, M., Mills, J., Allen, S., & Tee, S. (2020). Managing the mutations: Academic misconduct in Australia, New Zealand and the UK. *International Journal for Educational Integrity*, 16(1), 6. <https://doi.org/10.1007/s40979-020-00055-5>
- Blankenberger, B., & Williams, A. M. (2020). COVID and the impact on higher education: The essential role of integrity and accountability. *Administrative Theory & Praxis*, 42(3), 404–423. <https://doi.org/10.1080/10841806.2020.1771907>
- Boitshwarelo, B., Reedy, A. K., & Billany, T. (2017). Envisioning the use of online tests in assessing twenty-first century learning: A literature review. *Research and Practice in Technology Enhanced Learning*, 12(1), 16. <https://doi.org/10.1186/s41039-017-0055-7>

- Bretag, T., Harper, R., Burton, M., Ellis, C., Newton, P., Rozenberg, P., Saddiqui, S., & van Haeringen, K. (2019). Contract cheating: A survey of Australian university students. *Studies in Higher Education, 44*(11), 1837–1856. <https://doi.org/10.1080/03075079.2018.1462788>
- Bretag, T., Harper, R., Burton, M., Ellis, C., Newton, P., van Haeringen, K., Saddiqui, S., & Rozenberg, P. (2019). Contract cheating and assessment design: Exploring the relationship. *Assessment & Evaluation in Higher Education, 44*(5), 676–691. <https://doi.org/10.1080/02602938.2018.1527892>
- Bryant, P., & Ruello, J. (2019). One system to examine them all: Defining the complexities of implementing an institution wide online exam model. In Y. W. Chew, K. M. Chan, & A. Alphonso (Eds.), *Personalised Learning. Diverse Goals. One Heart. ASCILITE 2019 Singapore* (pp. 370–374).
- Butler-Henderson, K., & Crawford, J. (2020). A systematic review of online examinations: A pedagogical innovation for scalable authentication and integrity. *Computers & Education, 159*, 104–124. <https://doi.org/10.1016/j.compedu.2020.104024>
- Coghlan, S., Miller, T., & Paterson, J. (2021). Good Proctor or “Big Brother”? Ethics of Online Exam Supervision Technologies. *Philosophy & Technology, 34*(4), 1581–1606. <https://doi.org/10.1007/s13347-021-00476-1>
- Cramp, J., Medlin, J., Lake, P., Sharpe, C., Lake, P., University of South Australia, Sharp, C., & University of South Australia. (2019). Lessons learned from implementing remotely invigilated online exams. *Journal of University Teaching and Learning Practice, 16*(1), 137–155. <https://doi.org/10.53761/1.16.1.10>
- Curtis, G. J., Slade, C., Bretag, T., & McNeill, M. (2021). Developing and evaluating nationwide expert-delivered academic integrity workshops for the higher education sector in Australia. *Higher Education Research & Development, 1*–16. <https://doi.org/10.1080/07294360.2021.1872057>
- Dawson, P. (2020). *Defending Assessment Security in a Digital World: Preventing E-Cheating and Supporting Academic Integrity in Higher Education* (1st ed.). Routledge. <https://doi.org/10.4324/9780429324178>
- Dawson, P., & Henderson, M. (2017). How Does Technology Enable Scaling Up Assessment for Learning? In D. Carless, S. M. Bridges, C. K. Y. Chan, & R. Glofcheski (Eds.), *Scaling up Assessment for Learning in Higher Education* (pp. 209–222). Springer. https://doi.org/10.1007/978-981-10-3045-1_14
- Dawson, P., & Sutherland-Smith, W. (2018). Can markers detect contract cheating? Results from a pilot study. *Assessment & Evaluation in Higher Education, 43*(2), 286–293. <https://doi.org/10.1080/02602938.2017.1336746>
- Debus, J. C. W., & Lawley, M. (2016). Benefits and drawbacks of computer-based assessment and feedback systems: Student and educator perspectives: Benefits and drawbacks of computer-based feedback. *British Journal of Educational Technology, 47*(2), 294–301. <https://doi.org/10.1111/bjet.12232>
- Dendir, S., & Maxwell, R. S. (2020). Cheating in online courses: Evidence from online proctoring. *Computers in Human Behavior Reports, 2*, 100033. <https://doi.org/10.1016/j.chbr.2020.100033>
- Douglas, M., Wilson, J., & Ennis, S. (2012). Multiple-choice question tests: A convenient, flexible and effective learning tool? A case study. *Innovations in Education and Teaching International, 49*(2), 111–121. <https://doi.org/10.1080/14703297.2012.677596>

- Ellis, C., van Haeringen, K., Harper, R., Bretag, T., Zucker, I., McBride, S., Rozenberg, P., Newton, P., & Saddiqui, S. (2020). Does authentic assessment assure academic integrity? Evidence from contract cheating data. *Higher Education Research & Development, 39*(3), 454–469. <https://doi.org/10.1080/07294360.2019.1680956>
- Ellis, S., & Barber, J. (2016). Expanding and Personalising Feedback in Online Assessment: A Case Study in a School of Pharmacy. *Practitioner Research in Higher Education, 10*(1), 121–129.
- Fask, A., Englander, F., & Wang, Z. (2014). Do Online Exams Facilitate Cheating? An Experiment Designed to Separate Possible Cheating from the Effect of the Online Test Taking Environment. *Journal of Academic Ethics, 12*(2), 101–112. <https://doi.org/10.1007/s10805-014-9207-1>
- Gehringer, E., & Peddycord III, B. (2013). Experience with Online and Open-Web Exams. *Journal of Instructional Research, 2*, 10–18. <https://doi.org/10.9743/JIR.2013.2.12>
- Harper, R., Bretag, T., & Rundle, K. (2021). Detecting contract cheating: Examining the role of assessment type. *Higher Education Research & Development, 40*(2), 263–278. <https://doi.org/10.1080/07294360.2020.1724899>
- Holden, O. L., Norris, M. E., & Kuhlmeier, V. A. (2021). Academic Integrity in Online Assessment: A Research Review. *Frontiers in Education, 6*. <https://doi.org/10.3389/educ.2021.639814>
- Jaap, A., Dewar, A., Duncan, C., Fairhurst, K., Hope, D., & Kluth, D. (2021). Effect of remote online exam delivery on student experience and performance in applied knowledge tests. *BMC Medical Education, 21*(1), 86. <https://doi.org/10.1186/s12909-021-02521-1>
- James, R. (2016). Tertiary student attitudes to invigilated, online summative examinations. *International Journal of Educational Technology in Higher Education, 13*(1), 19. <https://doi.org/10.1186/s41239-016-0015-0>
- James, L. T., & Casidy, R. (2018). Authentic assessment in business education: Its effects on student satisfaction and promoting behaviour. *Studies in Higher Education, 43*(3), 401–415. <https://doi.org/10.1080/03075079.2016.1165659>
- Kharbat, F. F., & Abu Daabes, A. S. (2021). E-proctored exams during the COVID-19 pandemic: A close understanding. *Education and Information Technologies, 26*(6), 6589–6605. <https://doi.org/10.1007/s10639-021-10458-7>
- Kolski, T., & Weible, J. (2018). Examining the Relationship Between Student Test Anxiety and Webcam Based Exam Proctoring. *Online Journal of Distance Learning, 21*(3), 1–15.
- Ladyshewsky, R. K. (2015). Post-graduate student performance in ‘supervised in-class’ vs. ‘unsupervised online’ multiple choice tests: Implications for cheating and test security. *Assessment & Evaluation in Higher Education, 40*(7), 883–897. <https://doi.org/10.1080/02602938.2014.956683>
- Linden, K., & Gonzalez, P. (2021). Zoom invigilated exams: A protocol for rapid adoption to remote examinations. *British Journal of Educational Technology, 52*(4), 1323–1337. <https://doi.org/10.1111/bjet.13109>
- Lloyd, N., Sealey, R., & Logan, M. (2021). Balancing the COVID-19 Disruption to Undergraduate Learning and Assessment with an Academic Student Support Package: Implications for Student Achievement and Engagement. *Student Success, 12*(2), 61–71. <https://doi.org/10.5204/ssj.1933>

- Milone, A. S., Cortese, A. M., Balestrieri, R. L., & Pittenger, A. L. (2017). The impact of proctored online exams on the educational experience. *Currents in Pharmacy Teaching and Learning*, 9(1), 108–114. <https://doi.org/10.1016/j.cptl.2016.08.037>
- Moore, C. P. (2018). Adding authenticity to controlled conditions assessment: Introduction of an online, open book, essay based exam. *International Journal of Educational Technology in Higher Education*, 15(1), 26. <https://doi.org/10.1186/s41239-018-0108-z>
- Munoz, A., & Mackay, J. (2019). An online testing design choice typology towards cheating threat minimisation. *Journal of University Teaching & Learning Practice*, 16(3). <https://doi.org/10.53761/1.16.3.5>
- Myry, L., & 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), 119–132. <https://doi.org/10.1177/1469787415574053>
- Okada, A., Noguera, I., Alexieva, L., Rozeva, A., Kocdar, S., Brouns, F., Ladonlahti, T., Whitelock, D., & Guerrero-Roldán, A. (2019). Pedagogical approaches for e-assessment with authentication and authorship verification in Higher Education. *British Journal of Educational Technology*, 50(6), 3264–3282. <https://doi.org/10.1111/bjet.12733>
- Pitt, E., & Winstone, N. (2020). Towards Technology Enhanced Dialogic Feedback. In M. Bearman, P. Dawson, R. Ajjawi, J. Tai, & D. Boud (Eds.), *Re-imagining University Assessment in a Digital World* (pp. 79–94). Springer International Publishing. https://doi.org/10.1007/978-3-030-41956-1_7
- Rahman, A. (2021). Using Students' Experience to Derive Effectiveness of COVID-19-Lockdown-Induced Emergency Online Learning at Undergraduate Level: Evidence from Assam, India. *Higher Education for the Future*, 8(1), 71–89. <https://doi.org/10.1177/2347631120980549>
- Reedy, A., Pfitzner, D., Rook, L., & Ellis, L. (2021). Responding to the COVID-19 emergency: Student and academic staff perceptions of academic integrity in the transition to online exams at three Australian universities. *International Journal for Educational Integrity*, 17(1), Article 1. <https://doi.org/10.1007/s40979-021-00075-9>
- Rolim, C., & Isaias, P. (2019). Examining the use of e-assessment in higher education: Teachers and students' viewpoints. *British Journal of Educational Technology*, 50(4), 1785–1800. <https://doi.org/10.1111/bjet.12669>
- Stone, C. (2017). *Opportunity through online learning: Improving student, access, participation and success in higher education* (Guidelines for Improving Student Outcomes in Online Learning) [National Guidelines]. National Centre for Student Equity in Higher Education.
- Sullivan, D. P. (2016). An Integrated Approach to Pre-empt Cheating on Asynchronous, Objective, Online Assessments in Graduate Business Classes. *Online Learning*, 20(3), Article 3. <https://doi.org/10.24059/olj.v20i3.650>
- Villarroel, V., Bloxham, S., Bruna, D., Bruna, C., & Herrera-Seda, C. (2018). Authentic assessment: Creating a blueprint for course design. *Assessment & Evaluation in Higher Education*, 43(5), 840–854. <https://doi.org/10.1080/02602938.2017.1412396>
- Vos, L. (2015). Simulation games in business and marketing education: How educators assess student learning from simulations. *The International Journal of Management Education*, 13(1), 57–74. <https://doi.org/10.1016/j.ijme.2015.01.001>

Warwick Academic Development Centre (2022). Principles of Assessment Design. Principle 1 - Assessment tests intended learning outcomes. https://warwick.ac.uk/fac/cross_fac/academic-development/assessmentdesign/assessmentdesignprinciples/