

RESEARCH

Free and Open Access

Online assessment in higher education: a mapping review and narrative synthesis

Qian Liu ^{1*}, Anjin Hu ^{1,2} and Ben Daniel ¹

*Correspondence:
qian.liu@otago.ac.nz
Higher Education Development
Centre,
University of Otago,
65 Union Pl, Dunedin,
New Zealand, 9016
Full list of author information is
available at the end of the article

Abstract

Online assessment takes many forms. While there have been reviews on a particular online assessment approach (e.g., online examinations or tests), there has not been a knowledge synthesis that considers online assessment research as a whole. A holistic understanding of online assessment research is important, as it recognizes differences in assessment approaches and the role of technologies in assessment. This understanding helps researchers navigate the heterogeneous body of research and allows educators to make research-informed improvements. To establish such an understanding, we analyzed 235 articles, following a mapping review and a narrative synthesis method. The findings revealed tests, assignments and skills assessments are major online assessment approaches. While research into tests reported using online technologies mainly to substitute or augment existing assessment, research into assignments and skills assessments was more likely to report using online technologies to modify or redefine assessment. Further, we identified disparities across the three approaches regarding academic misconduct, assessment validity, and reliability. We also identified a dearth of comparative research and a reliance on overall satisfaction and short-term self-reported impact measures as indicators of success. We discuss the implications of this review to provide insights for institutions and educators seeking to improve online assessment practices.

Keywords: Online assessment, SAMR, Higher education, Mapping review, Narrative synthesis

Introduction

Technologies are pervasive in assessment practices in higher education. They have been used in various ways to support exam delivery, manage assignment submission, improve assessment feedback, and, not the least, store and communicate assessment results (Bearman et al., 2023). In the contemporary world where work and lives are increasingly mediated by technologies, they are also essential to assessment tasks that seek to prepare



© The Author(s). 2024 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

university graduates for future careers and lives (Geertshuis et al., 2022). The recent shift to remote delivery of teaching further rendered technologies high-profile, as university educators had to leverage the affordances of technologies to enable assessment in online environments (Lee & Fanguy, 2022).

Within the research literature, there has been ongoing effort dedicated to investigating the intersection between technology and assessment. This research was initially termed computer-assisted assessment (Conole & Warburton, 2005) and more recently as e-assessment or online assessment (St-Onge et al., 2022). The change in the use of terms reflects the progressive development and application of online technologies in assessment. Early explorations relied on the use of tailor-made computer programs to assess learning within specific disciplines or institutions (Fiddes et al., 2002). Recent developments involved the application of a breadth of online tools, platforms and resources to assessment practices, often across disciplines and institutions (Lee & Fanguy, 2022). Findings generally confirmed the potential of online technologies in enhancing or transforming assessment practices but at the same time identified challenges, including but not limited to the need to upskill staff in assessment design (Snodgrass et al., 2014), ensure technology accessibility and scalability (Avila et al., 2016), prepare students for online assessment (Brink & Lautenbach, 2011), authenticate learners (Levy et al., 2011) and prevent academic misconduct (Buckley & Cowap, 2013).

While empirical research abounds internationally, there have been limited knowledge syntheses on online assessment in higher education. Among them, Gikandi et al. (2011), through a review of 18 studies, concluded that online formative assessment provided additional learning opportunities and improved learner engagement, but cautioned the need for designing authentic and engaging assessment tasks. Boitshwarelo et al. (2017), through a review of 50 studies, suggested that online tests can be appropriate for assessing twenty-first century learning when they are used formatively and in combination with other types of assessment. A recent review of 36 studies further suggested that while online examinations could support student learning and performance and reduce staff grading workload, they need to address issues relating to accessibility, validity, reliability and potential cheating during implementation (Butler-Henderson & Crawford, 2020).

The above reviews allowed us to gauge the efficacies of and challenges associated with a particular online assessment approach. They are, however, focused on one assessment function (e.g., formative assessment) or task (e.g., online tests or examinations), which is restricted in scope and may overlook the differences and commonalities among different online assessment approaches. A comprehensive review would capture this diverse body of research and establish a more holistic understanding of current research into online assessment. Such a review would also help educators understand the current landscape of

online assessment research and become informed and versatile as they seek to redesign and transform assessment through online technologies.

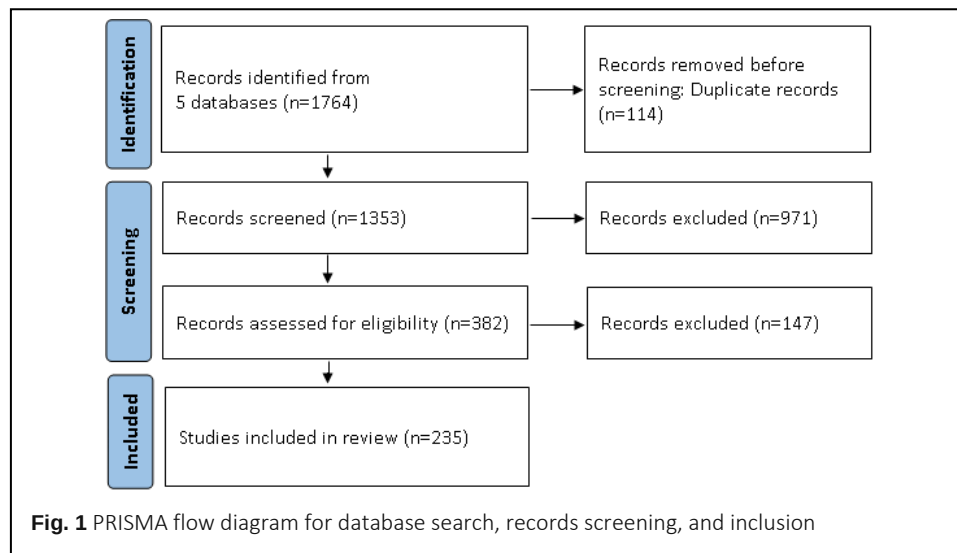
To this end, we conducted a comprehensive review of online assessment in higher education. The review aimed to establish the current understanding of online assessment and expose knowledge gaps. Three research questions (RQs) guided the review: what online assessment approaches have been reported in the research literature (RQ1); what is the role of technologies in online assessment (RQ2), and; what is the current understanding concerning different assessment approaches and different roles technologies serve (RQ3).

For this review, we define online technologies as digital devices, tools and platforms that connect users to the internet. We regard assessment broadly as the process of making a judgement about student work or performance, which can be derived from the student, peers, teachers or machine (Bearman et al., 2023). We accept the notion of assessment of learning, which emphasizes the summative purpose of assessment as a tool for measurement and accreditation, and the notion of assessment for learning, which emphasizes the formative role of assessment in promoting learning (Wiliam, 2011). However, we also align with Boud et al. (2018), acknowledging that a clear boundary between formative and summative assessment functions may be hard to observe in contemporary assessment practices. Thus, our orientation towards online assessment is inherently broad: online assessment refers to the assessment that takes place in the online environment or uses online technologies during the assessment process. Importantly, we do not restrict online assessment to online or distance learning modalities. While fully online or distance courses necessitate the use of online assessment, other modalities such as web-enhanced and blended learning have also adopted online assessment practices. Our review therefore considers online assessment across different learning modalities.

Method

Article search and screening

The review procedure (Figure 1) followed the PRISMA guideline (Page et al., 2021). During the database search phase, we formulated the following search string through consultations with an experienced librarian: ((internet* OR online* OR web* OR electronic* OR computer* OR digital* OR eportfolio) AND (assess* OR evaluat* OR exam* OR test* OR assignment*)). We applied the search string to five databases in June 2021, including Web of Science, ERIC, ProQuest Education, Academic Search Complete and Education Research Complete. We further limited the search to full-text, peer-reviewed articles published since 2011, written in English and in the higher education



subject area. This led to 1764 records being identified and 1353 records after duplication removal.

Two reviewers (QL and AH) conducted title and abstract screening and full-text screening independently. Articles were included if they were (1) empirical studies, (2) in the higher education context, and (3) focused on online assessment. Articles were excluded if they were (1) reviews, theoretical or opinion pieces, (2) outside of higher education, (3) not on assessment (e.g., online courses or learning tools), or (4) not on online technologies (e.g., assessment in general). Disagreement between the reviewers was discussed to reach a consensus. Given that the included articles adopted diverse research designs and our review, unlike systematic reviews, did not serve to inform interventions in practice, we conducted quality appraisal by way of ensuring that included articles reported research aims, questions or methodological details. This is in line with the recent quality appraisal guide that recommends using quality appraisal measures to understand the quality of research findings rather than to screen studies (Hong et al, 2018). In total, 382 articles were included after title and abstract screening, and 235 included after full-text screening and quality appraisal.

The reviewers then extracted data to a spreadsheet. Information extracted included first author, year of publication, journal, research aim or question, methods, findings, and conclusion. Overall, the reviewed articles ($N=235$) covered 101 peer-reviewed journals, with *Computers & Education* ($n=20$), *Assessment & Evaluation in Higher Education* ($n=11$) and *BMC Medical Education* ($n=10$) contributing the most. The articles captured research from 39 unique countries and regions, with the USA ($n=47$), Australia ($n=28$), the UK ($n=24$), Taiwan ($n=20$), Mainland China ($n=12$), Spain ($n=11$) and Germany ($n=10$) dominating the sample. Various research designs have been employed, including,

for instance, questionnaires, quasi-experiments, pre- and post-tests, interviews, focus groups, case studies, and design-based methods.

Analysis and synthesis

We combined mapping review and narrative synthesis methods to answer the stated RQs. Mapping review focuses on categorizing studies on a given topic to enable contextualized knowledge synthesis (Grant & Booth, 2009). We used this method to first inductively identify major online assessment approaches reported in the literature. For example, studies on the use of formative quizzes with multiple-choice questions were coded as ‘formative quizzes’ and grouped together with studies using quizzes with drop-down menu questions and short answer questions. Studies coded as ‘formative quizzes’, ‘diagnostic tests’, and ‘examinations’ were further grouped together as ‘tests’, which was identified as an assessment approach.

We then drew on the SAMR (Substitution, Augmentation, Modification and Redefinition) model to further map out variations in technology use across assessment approaches. The SAMR model provides a useful understanding of technology use in education, and has been widely applied to examine the role of technology in pedagogical activities (Crompton & Burke, 2020). The model specifies four levels of technology use, from the lowest level of substitution to augmentation, modification and eventually redefinition (Hamilton et al., 2016). Substitution refers to using technology as a direct tool to substitute existing practices without functional improvement. Augmentation refers to using technology to substitute existing practices with functional improvement. Modification refers to the use of technology for significant task redesign, and redefinition refers to the use of technology to create new tasks or practices previously inconceivable. In this review, we adapted the definitions of the four levels of technology use so that they are relevant to online assessment (see Table 1) and used the model to map out the role of technologies in online assessment.

Table 1 Adapted definition of substitution augmentation modification and redefinition in relation to online assessment

Categories	Definition adapted to online assessment
Redefinition	Digital technology enables the new forms of assessment practice that could otherwise not be able to realise.
Modification	Digital technology initiates redesign or rearrangement of aspects of existing assessment practice.
Augmentation	Digital technology improves aspects of the existing assessment practice incrementally.
Substitution	Digital technology re-produces the existing assessment practice that can be achieved without the digital technology.

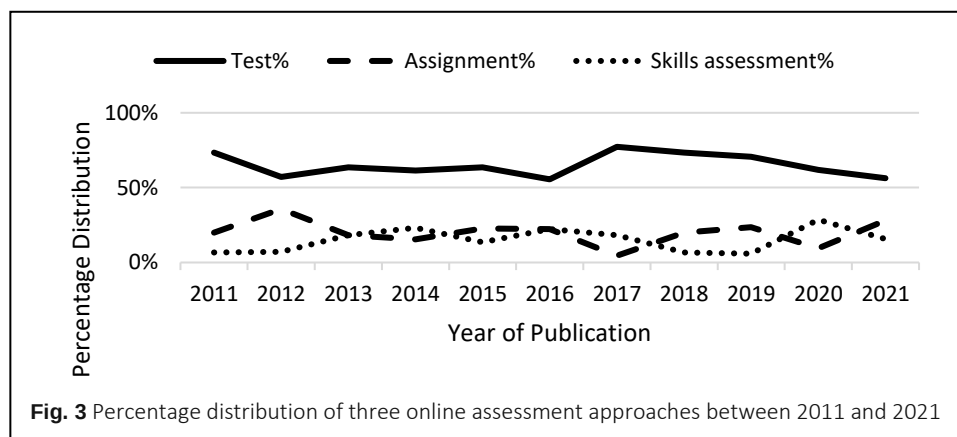
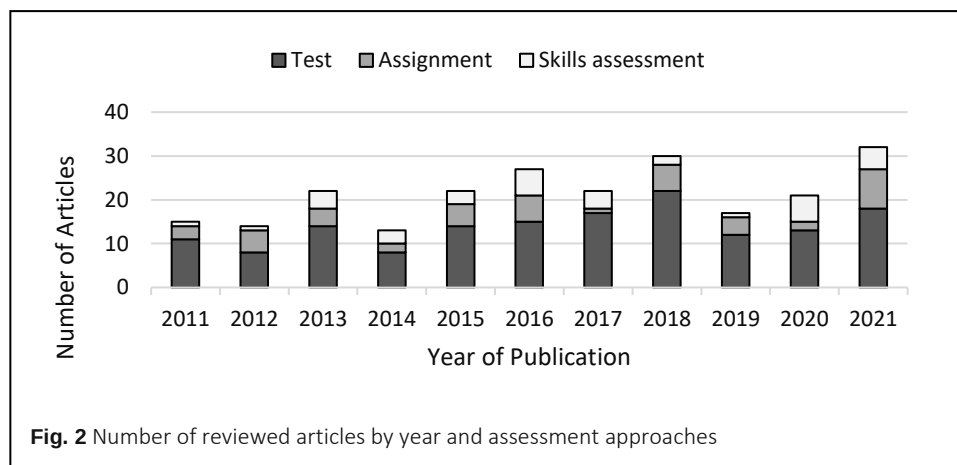
Once we categorized online assessment approaches and the variations in technology use, we used the narrative synthesis method (Popay et al., 2006) to synthesize research findings within an assessment approach and at a particular SAMR level. This was an iterative process involving textual analysis for synthesizing and comparing research findings within and across different online assessment approaches and ways of technology use.

Result

Mapping of current research

There has been a gradual increase in research (Figure 2). The first half of the review period (2011–2015) contributed an average of 17 articles per year, and the second half contributed an average of 24 per year.

Three major online assessment approaches were identified in this literature, which were (1) tests; (2) assignments, and; (3) skills assessments. Despite the overall increase in research, tests remained the major online assessment approach (65%), followed by assignments (20%), and skills assessment (15%) throughout the review period (Figure 3).



Online tests

Tests (n=152) were the most common online assessment approach, which focused on assessing learners' lower level of cognition, such as recall or comprehension (Kratwohl, 2002), through the use of online technologies, often in the form of 'quizzes', 'tests' or 'examinations'. This approach typically involved students responding to structured, objective or semi-objective questions. Multiple-choice question was widely adopted within this approach (Wilson et al., 2011), although there were reported use of more sophisticated question types such as long drop-down menu questions (Huwendiek et al., 2017) or open-ended questions (Zlatovic et al., 2015).

There were considerable variations in what the assessment served and the way the assessment was implemented. Quizzes were implemented regularly during learning to help students self-assess their understanding, serving for formative assessment purposes (Thomas et al., 2017). Tests were used for both summative and formative purposes during the course of learning (e.g., pre-learning tests, or mid-term tests): they served to diagnose learners' existing knowledge prior to learning (Carr et al., 2017), were implemented during learning to provide feedback (Balta et al., 2018), or contributed to summative assessment (Bausili, 2018). Examinations on the other hand often served for summative purposes that directly contributed to learners' final grades and performance evaluation (Daffin & Jones, 2018).

Online assignments

Assignments (n=47) involved assigning students specific tasks and requiring them to present the task outcome later. The tasks involved, for instance, essay writing (Noroozi et al., 2016), journalling (Zhan, 2021), individual or group projects (Alden, 2011), laboratory work (Whitworth & Wright, 2015), and digital production (Davies et al., 2011). Task outcomes included written pieces (i.e., essays and reports) or digital artefacts (i.e., websites or videos), and were used for self-assessment (Hwang et al., 2015), peer-assessment (Lin, 2019) or teacher-led assessment (Gray et al., 2012). Unlike tests, assignments usually took an extended period of time and were fully embedded in the process of student learning: they had a clear developmental emphasis, encouraging students to develop higher level cognitive skills and capabilities (e.g., application, analysis, evaluation and creation) through ongoing engagement with the assessment task, learning resources and, in many cases, their peers.

Although various online technologies were used to facilitate assignment tasks, students were likely to switch between online and offline. For instance, some studies (Fernando, 2018; Noroozi et al., 2016) reported having students write essays, which only involved using a few databases and a word-processing package but subsequently introduced additional online platforms for peer assessment.

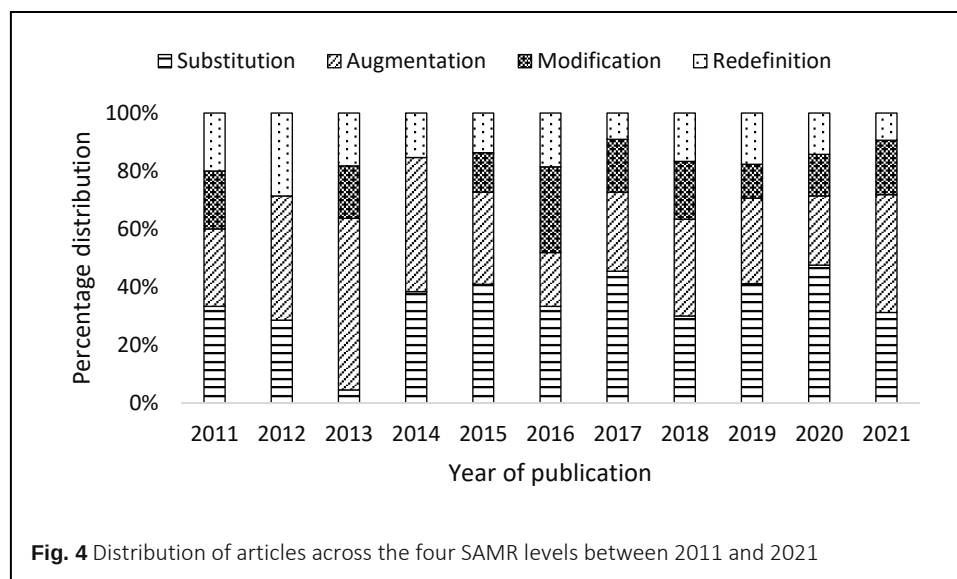
Online skills assessments

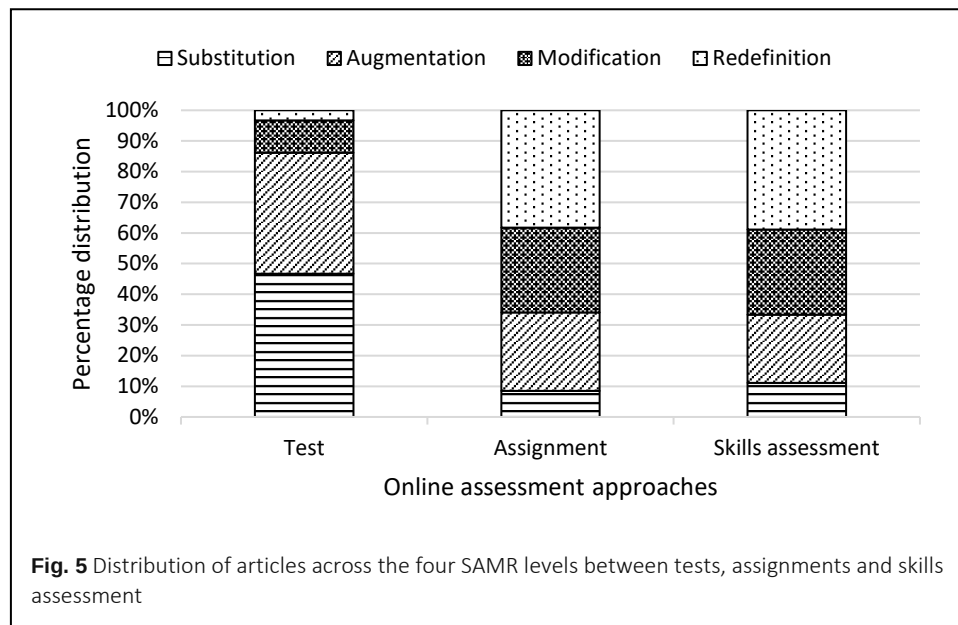
A further approach focused directly on assessing student performance in practice, which involved the application of multiple skills at different levels ($n=36$). Given the diversity of skills being assessed, the assessment practices varied considerably so did online technologies being used to support the assessment. In the most straight-forward form, online technologies were used to develop electronic rubrics to help assess, for instance, second language skills or professional attitudes (Haack et al., 2017). More complex technology use involved assessing student performance with higher level of authenticity through simulation (Craft & Ainscough, 2015), using video assessments to capture student performance (Hay et al., 2013), and enabling portfolio-based assessments to integrate aspects of student development, performance and reflection (Tinoca & Oliveira, 2013).

Variations in the role of online technologies

Regarding the differences in the way online technologies were used, around 34% of studies were assigned at the ‘substitution’ level, 34% were at ‘augmentation’ level, 17% were at the ‘modification’ level and 16% were at the ‘redefinition’ level. Throughout the review period, studies assigned at the substitution and augmentation levels dominated the sample (Figure 4).

Our mapping of the studies against the SAMR model further showed variations in technology use across assessment approaches (Figure 5). For tests, 47% of studies used online technologies to substitute existing assessments, 39% used online technologies to augment existing assessments, 11% used online technologies for modification purposes,





and 3% used online technologies to redefine assessments. For assignments, 9% of studies used online technologies to substitute assessments, 26% used online technologies to augment existing assessments, 28% used online technologies for modification purposes, and 38% sought to redefine assessments through online technologies. For skill assessments, 11% of studies used online technologies to substitute existing assessments, 22% used online technologies to augment existing studies, 28% used online technologies for modification purposes, and 39% sought to redefine assessments through online technologies.

The mapping analysis showed that while most studies focused on using online technologies to substitute or augment existing assessments, the functions that online technologies served depended on the assessment approach. While tests predominantly integrated online technologies at lower levels (i.e., substitution or augmentation), assignments and skill assessments focused more on online technologies' modification and redefinition potential.

Synthesis of current research

This section describes the result from a narrative synthesis of current research into each of the three online assessment approaches. Within each approach, studies were organized according to their assigned SAMR level to expose similarities and differences in technology use.

Current research into online tests

Tests at the substitution level. Most studies on using online technologies for tests were categorized at the substitution level (n=71). Technologies were used to deliver existing paper-based tests online. Students were attitudinally mostly positive towards online tests because of the efficiency in assessment delivery (Deutsch et al., 2012). However, there were concerns regarding data security and calls for measures in data management and user authentication (Brink & Lautenbach, 2011). Studies also captured issues relating to fairness in online tests (Liu et al., 2015), suggesting that online tests may disadvantage learners who had low self-efficacy in using technologies, especially if the learners had never sat online tests before (Daniels & Gierl, 2017).

Although none of the studies focused on measuring the effectiveness of institutional interventions that better prepare learners for online tests, there was a consensus that effective delivery of online test that substitutes existing paper-based test required training for teachers and for students (Brink & Lautenbach, 2011). Planning, piloting, and providing ongoing technical support were recommended as importance means for successful substitution (Khalaf et al., 2020). Researchers also called for assessment policies to be sensitive to online modalities (Liu et al., 2015).

Many studies sought to compare the effectiveness of online and paper-based tests in assessing student performance and explored a related issue, cheating. The findings, however, were less consistent. Rivera-Mata (2021) reported that cheating during a test was related to the type of class and students rather than the face-to-face or online modalities. Others, however, reported differences in test scores between online and paper-based tests, especially when the online test was not proctored (Alessio et al., 2018), suggesting that cheating was a concern for online tests. Researchers have also identified interventions that reduced cheating during online tests, which included implementing browser lockdown and time controls during tests, designing complex questions with multiple variants, and introducing remote proctoring or artificial intelligence-based proctoring (Paredes et al., 2021).

Tests at the augmentation level. Studies at the augmentation level (n=60) employed online tests to gain better understanding of learners' current knowledge and provide additional learning opportunities. There was an emphasis on using online technologies to enable timely and often automated feedback (Herbert et al., 2019) or design tests that incorporated multimedia resources or sophisticated types of questions (Cerutti et al., 2019). These studies unanimously reported student appreciation of formative feedback from online tests (Chen & Chuang, 2012). Overall, online tests that served to augment learning allowed learners to self-assess their knowledge (Mina et al., 2011) and monitor their progress (Bayrak, 2021), which led to better academic performance later on (Swart & Meda, 2021).

Teachers were similarly appreciative of online tests primarily because they allowed frequent formative feedback to large groups of students (Balter et al., 2013). Online tests were reported to reduce marking workload through automated grading of objective (Figueroa-Canas & Sancho-Vinuesa, 2021) and open-ended questions (Nehm & Haertig, 2012).

Tests at the modification level. Studies at the modification level ($n=16$) used online technologies to fully embed tests in the learning process creating responsive (Yasuda et al., 2021) or dynamic (Ebadi & Rahimi, 2019) learning environments where learning progress was based on student responses to previous test items. These tests were interactive, individually customized (Kamrood et al., 2019). They were popular among learners (Yang & Qian, 2017), associated with better academic outcomes (Pezzino, 2018), and were able to guide students' further learning (Ebadi et al., 2018).

Tests at the redefinition level. Redefining tests through online technologies comprised a small group of studies ($n=5$). These studies used online technologies to create highly innovative test experiences for students. There were instances of having students create test questions rather than simply sitting in tests (Yu & Wu, 2016), using scrambled text to assess reading comprehension rather than mastery of vocabulary (Thompson & Braude, 2016), using knowledge mapping to promote holistic understanding (Ho et al., 2018), and using social network analysis to influence peer feedback-seeking (Lin & Lai, 2013). Although the studies designed online tests differently, they had in common an emphasis on using technologies to enable active and interactive learning that led to better learner engagement. The studies reported students' attitudinal acceptance towards the test and improved academic outcomes. However, given that each study took a unique approach to assessment design, it was premature to synthesize more nuanced research findings across the studies.

Current research into online assignments

Assignments at the substitution level. Studies categorized as using online assignment for substitution ($n=4$) used online technologies to deliver existing assignments and focused on comparing the efficacy between online and traditional deliveries. The assignments were all in written formats, and online technologies were used to assist with writing and composition (Cheung, 2016), marking and feedback (Grieve et al., 2016), or peer assessment (Lin, 2019). One study reported that writing assignments online improved the quality of student work (Cheung, 2016). Other studies, however, showed that students might not be attitudinally more positive towards online marking and feedback (Grieve et al., 2016), nor would they achieve better outcomes through online peer assessment of assignments (Lin, 2019). It seemed that the modality effect (efficacy of online versus traditional delivery) was associated with issues such as assessment

design or perceived social presence rather than the online technology. Another study that investigated student choice of online assignments further reported that while students had opportunities to be more innovative, the majority chose traditional written assignments (Flavin, 2021).

Assignments at the augmentation level. Studies categorized as using online assignments for augmentation (n=12) also employed written assignments, or in some cases, laboratory or project work where written reports were part of the assessed project outcome. Online technologies were used to improve the marking and feedback process, creating formative learning during assignment completion.

There were differences in marks between technology-enabled marking (e.g., statement banks or automated marking) and teacher-led marking, and technology-enabled marking was mainly praised for efficiency and timely feedback (Reilly et al., 2014). Integrating online technologies also enabled students' self-assessment and peer sharing, helping students develop generic cognitive and writing skills (Hwang et al., 2015). However, the efficacy of online assignments varied by individuals: students behaved differently during assignment, and their task interpretation, goal setting, perceived autonomy, interactivity and interests were similarly different, which were associated with their assignment performance (Beckman et al., 2021).

Assignments at the modification level. Studies that sought to modify assignment tasks (n=13) used online technologies to enable peer feedback for written assignments (Chew et al., 2016) or peer assessment of student contribution in group work (Alden, 2011). These assignments involved significant redesign of the assessment procedures. Findings suggested that enabling peer assessment improved assignment quality (Mostert & Snowball, 2013) and course completion (Formanek et al., 2017), and developed writing, critical thinking and reflection skills (Zheng et al., 2018). There were, however, different views regarding peer assessment's reliability and validity, suggesting that training before peer assessment being important for feedback quality (Chew et al., 2016).

Assignments at the redefinition level. Studies that sought to redefine assignments through online technologies (n=18) took two different approaches. The first used online technologies to help students prepare for assignments rather than using them to assess written outputs (Fernando, 2018). Students engaged in ongoing collaborative writing activities in online spaces (Alvarez et al., 2012), received regular feedback from peers and used a range of multimedia resources to prepare for components of an assignment (Fernando, 2018). There were also opportunities for students to use text-matching tools or discussion forums to improve their work and avoid academic misconduct (Buckley & Cowap, 2013). This way of technology use contributed to knowledge construction, enabled learner autonomy, facilitated reflection and critical thinking, and reduced the plagiarism ratio (Alvarez et al., 2012).

The second approach focused on digital production, where students produced digital artefacts rather than compiling written reports as the assignment output. Assignments following this approach engaged students in creating, for instance, films (Cheng & Hou, 2015), videos (Fang et al., 2021), programming scripts (Wang, 2019), and online wikis (Davies et al., 2011). They engaged students in active learning (Fang et al., 2021), facilitated deep approaches to learning (Cheng & Tsai, 2012), and developed digital skills that cannot be learnt or assessed without the use of online technologies (Nielsen et al., 2020). However, researchers highlighted the importance of pedagogy-informed design and implementation in preparing students to learn from the assignment projects (Cheng & Hou, 2015; Wang, 2019).

Current research into online skills assessments

Skills assessments at the substitution level. Studies that used online technologies to substitute existing skills assessments (n=4) were typically focused on digitalizing the existing assessment rubrics (Haack et al., 2017) or the assessment delivery (Snodgrass et al., 2014) in clinical or healthcare settings. Findings suggested that the digitalization of existing skills assessment was not inferior to traditional delivery (Kaliyadan et al., 2014) and enabled immediate and customized feedback (Haack et al., 2017). However, training was necessary for the implementation of digitalized skills assessment (Snodgrass et al., 2014).

Skills assessments at the augmentation level. Skills assessments at the augmentation level (n=8) differed from the previous group in that they explicitly sought to provide more formative learning opportunities through digitalization of existing skills assessment. These assessments provided additional opportunities for automated, self and peer assessments of a range of work-related skills, thus allowing students to learn from the assessment and demonstrate performance improvements (Ros et al., 2021). However, studies that focused on assessing technical skills in the clinical setting reported differences between automated assessment and clinician assessment, suggesting that automated assessment was best used for formative rather than summative purposes (Abdalla et al., 2020). In addition, one study noted the difference between students' self-perceived developments and actual performance developments, highlighting again the importance of feedback and guidance during the assessment activity (Lim et al., 2020).

Skills assessments at the modification level. Studies that modified skills assessments (n=11) were characterized as using online technologies to implement video assessment or to develop simulation-based assessment. Video assessment involved having students create videos to capture their performance in authentic learning or work environments at a distance and often asynchronously (Hay et al., 2013). Students were often allowed to practise multiple times, critique each other's performance and present their best

performance (Lai et al., 2020), thus effectively transforming the assessment activity from focusing on measurement of achievement to a demonstration and exhibition of achievement (Geertshuis et al., 2022). Video assessment was an effective assessment solution to distance programs or distributed learning environments (Hay et al., 2013). In addition, it was reported as an appropriate means to developing and assessing skills-based performances especially when it is accompanied by peer assessment and feedback (Hensiek et al., 2016).

Simulation-based assessment, on the other hand, used online technologies to create assessment that simulated authentic workplace scenarios or problems. The studies reported different simulation solutions with varying levels of fidelity but had in common in using simulation or simulated cases to assess complex work-related skills (e.g., decision-making and problem-solving) that are difficult to assess in a traditional university setting (Way et al., 2021). The findings suggested that simulation-based assessment allowed students to apply what they learnt and, to some extent, allowed educators to assess work-related skills in a university setting (Craft & Ainscough, 2015).

Skills assessments at the redefinition level. Studies at the redefinition level ($n=18$) focused on portfolio-based assessment through which learner progression was captured and achievement evaluated. Some studies asked students to create portfolios themselves (Bleasel et al., 2016); others used workplace performance data (e.g., students' clinical performance data in a healthcare system) as the performance portfolio (Sebok-Syer et al., 2019). These studies could all be regarded as taking a programmatic approach to assessment, measuring and supporting the ongoing development of knowledge, skills and competence over an extended period of learning. Additionally, many portfolio-based assessments had clear milestones and performance or task expectations, which created multiple formative feedback opportunities (Marinho et al., 2021). Studies reported that this assessment method created learner autonomy and led to meaningful reflections (Marinho et al., 2021), which motivated learners, facilitated self-directed learning and promoted professional development (Tinoca & Oliveira, 2013). However, researchers also noted that online tools or platforms used for portfolio-based assessment could be relatively complex, which might be a barrier for adoption if technical training and support was inadequate (Avila et al., 2016).

Discussion

To establish an overall understanding of online assessment research, which takes into consideration of different assessment approaches and different roles online technologies serve, we set out a knowledge synthesis. The 235 articles included ensured that our review was grounded in an evidence base that captured the heterogeneous body of research. In the paragraphs below, we discuss the review results in light of the current

research patterns, gaps for further research, and implications for online assessment practice.

Tests as the dominant online assessment approach

Our analysis identified three major online assessment approaches. Among them, tests have been the dominant one, with nearly two thirds of articles being focused on this assessment approach. Studies on online assignments and skills assessments have been comparatively limited. This result suggests that, despite the gradual increase in research over time (Figure 2), tests are still the most common assessment practice in contemporary higher education (Boitshwarelo et al., 2017) and researchers have been using online technologies mainly to support this assessment. Noticeably, the move towards assessment for learning in higher education was reflected in the use of online tests. The tests served to not only evaluate student outcomes at the end of a semester but also to capture students' current understanding and allow students to practise learning and receive feedback on their developments. However, if we accept there is an association between assessment formats and assessed learning outcomes (Biggs, 1996) then the reliance on online tests would suggest that online assessment in higher education remains likely to be focused on cognitive domain (Krathwohl, 2002) and restricted to lower-order cognition (Liu et al., 2023). Online assignments and skills assessment that focused on higher-order cognition or learner's competence and performance beyond cognitive understanding contributed, by contrast, consistently around 30% of included articles with no evidence of increase over the review period. This suggests that, while more and more online assessment tools and platforms have been made available, a paradigm shift in terms of what intended outcomes are assessed remains yet to occur.

Technology use across online assessment approaches

Relatedly, results from the SAMR categorisation showed that, regardless of the assessment approach, the vast majority of research reported using online technologies at levels of substitution and augmentation (Figure 5), and this pattern remained unchanged throughout the review period (Figure 4). They suggest that there has not been a greater use of more sophisticated online assessment across the higher education sector although online technologies have become more and more prevalent and accessible over the years. This finding is consistent with a recent analysis on the role of technology in transforming disciplinary education (Grainger et al., 2024). It points to the difference between using technology to enhance existing pedagogical or assessment design and using technology to drive new (or transform) pedagogy or assessment. There have been individual studies that explored learning experiences and outcomes that are unique to a specific online

assessment task (Calderon & Sood, 2020). However, such thinking has not become widespread in the research literature.

There were differences in technology use across three identified approaches. More than 85% of studies on online tests used technologies to substitute or augment existing assessment. This is in contrast to research into online assignments and skills assessment, each of which had around 60% of studies on using technologies for substitution or augmentation. One possible explanation for the difference in technology use might be that tests have traditionally been used for high-stakes assessment and have been widely used for assessment in large classes. In either case, innovative technology use (e.g., modification or redefinition) in an online test would be high-risk for educators and would be regulated by strict assessment policies. Instead, replication and augmentation use of technology that draws on established practices is less effortful and likely to comply with existing policy prescriptions and curriculum arrangements, hence more likely to be adopted widely.

There were reported additional advantages of technology use at levels of modification and redefinition. While technology improved the efficiency of online assessment delivery (e.g., less time and workload) across four SAMR levels, studies at modification and redefinition levels further reported technology as enabling learner engagement with each other and with the assessment task itself. This seems to suggest that the advantage of technology lies in the way it was used pedagogically rather than being inherently to the technology itself.

Areas for future research

Our review and analysis above further identified areas of disparity in current research, which we discuss here with an aim to inform future research. One apparent disparity across the three online assessment approaches relates to academic misconduct. Academic misconduct refers to a range of unethical behavior (e.g., cheating or plagiarism) some students engage while completing academic work (Christensen Hughes & McCabe, 2006). Many reviewed studies explored academic misconduct by comparing student performance between online and offline modalities. This reflects an assumption that academic misconduct is mainly related to students and triggered by specific technological affordances. The institutional interventions therefore focused on regulating student behavior through policies and technical procedures (Paredes et al., 2021). None of the studies referred to the appropriate use of student work produced during online assessment by educators and institutions, despite increased instances of misconduct in this regard (Binder et al., 2016). In addition, student academic misconduct was explored in online tests, occasionally in online assignments (Buckley & Cowap, 2013), but not in online skills assessments. It is unlikely that certain assessment approaches are completely

immune to academic misconduct, given that academic misconduct has become prevalent (Christensen Hughes & McCabe, 2006) and takes many forms (e.g., impersonation). More research is perhaps needed to investigate academic misconduct beyond cheating during online tests.

Assessment validity refers to the extent that an assessment measures what it is designed to measure, and reliability refers to the extent that an measurement is consistent and accurate (Darr, 2023). Across the three approaches, validity and reliability issues were investigated primarily at the substitution level. This suggests that researchers have spent considerable effort examining how replicating an existing assessment in the online environment influences assessment validity and reliability. Significantly less research has been conducted into understanding whether innovative online assessment practices at modification and redefinition levels are valid and reliable. Admittedly, innovative assessment practices may inherently align more with the notion of assessment for learning (Wiliam, 2011), therefore emphasizing the developmental function over the measurement function of assessment. However, establishing the validity and reliability of online assessment at modification and redefinition levels serves to ensure the worth of these innovative practices, which in turn facilitates further adoption and dissemination. More evidence on the efficacy of online assessment at higher levels of SAMR is therefore needed.

Methodologically, the studies reviewed often included measures of attitudinal responses to an online assessment. The result has consistently showed that, regardless of the online assessment and its SAMR level, students and teachers generally find it satisfactory and acceptable, especially when training and ongoing support are provided. Therefore, research that describes the overall acceptance of and satisfaction with online assessment may have reached maturity. Future research should include additional measures other than attitudinal response or seek to establish whether certain ways of designing and delivering a particular online assessment are likely to lead to superior satisfactory experiences than other alternatives.

In terms of the impact of online assessment, ample research has included self-reported measures on the development of knowledge and generic and disciplinary-specific skills. In many cases, these measures were obtained right after the assessment activity and thus could be regarded as immediate impact measures. With a few exceptions (e.g., portfolio-based assessments or formative quizzes in relation to final exam scores), the impact of an online assessment on students' long term academic performance has been infrequently reported. Future research could strengthen the research base on the impact of online assessment by including objective measures collected at the end of learning or after the learning.

Finally, although many studies at the substitutional level compared the learner performance in online assessment with that in traditional assessment, there was a dearth of comparative studies. Current findings on satisfaction and immediate self-reported learning gains have portrayed online assessment as encompassing multiple promising practices, each of which has been effective. Comparative studies would have been valuable in exposing essential assessment design elements that could lead to better student learning and performance across different educational contexts.

Limitations

The SAMR model has been criticized for lack of clarity between levels of technology use. We sought to mitigate this through adapting definitions of the four SAMR levels to the context of online assessment ahead of analysis. We further reviewed and revised the mapping result to ensure our mapping was based on collective agreement.

The database search underpinning this review was conducted in mid-2021, which meant that articles published since the second half of 2021 were not captured. We are nonetheless confident that the review provided an overall understanding of different assessment approaches and different ways of technology use in online assessment. However, research in online assessment is heterogenous and fast growing. The application of artificial intelligence (e.g., ChatGPT) in higher education, for instance, has already stirred much attention. This development was not captured in our review but is becoming an important research topic and is influencing higher education assessment practices.

We were unable to synthesize findings according to student demographics (e.g., first year undergraduates, or non-traditional learners). In reality, students differ in their digital skills and are likely to experience online assessment differently. Additionally, various technologies have been used for online assessment. While technology accessibility was not identified as a common challenge, there were occasional references to technology being a barrier to assessment. These issues fall outside of the scope of our review but are worthy of further analysis by future reviews.

Implications for universities and educators

Our review suggests that successful implementation of online assessment requires supportive institutional policies and procedures, effective training, and time and opportunities for educators and students to understand the assessment as well as familiarize with the technological artefacts. Universities that seek to adopt online assessment should therefore carefully plan for the move to online assessment as institutional-wide change initiative rather than as implementing a digital tool as an add-on

to the existing assessment processes. This certainly has significant resource, staff development, student onboarding, and curriculum change implications.

For educators, our review identifies avenues for transforming assessment practices through online technologies. Looking across the three approaches, we note that substitution was characterised as using technologies to digitalize assessment; augmentation was characterised as using technologies for additional feedback; modification was characterised as using technologies to mediate the process by which students prepare for, participate in and receive feedback from the assessment in a more interactive and effective manner, and finally; redefinition was characterised as using technologies for assessment that was often student-led and collaborative, aiming to develop holistic understanding and digital skills and to document the development of performance and achievement over time. Thus, the progressive transformation of assessment led by online technologies could be interpreted as the process by which technologies serve from being the tools for assessment delivery to becoming fully embedded in assessment design, mediating assessment tasks, feedback processes and measurement of achievement. At the same time, students are afforded more opportunities to engage with the whole assessment process, from those being assessed and receiving feedback to becoming involved in responding to feedback, assessing themselves and others, designing assessment activities, and documenting and exhibiting individual or collaborative performance during learning. This provides many opportunities for educators to change their assessment practices through online technologies.

Conclusions

This review captured current understanding of online assessment, identifying major online assessment approaches and the different functions online technologies served. While tests have been the dominate approach, in which technologies substituted existing assessment, online assignments and skills assessments involved more innovative assessment practices. There were disparities in current understanding relating to academic misconduct, assessment validity and reliability across the three approaches, and future research should adopt a comparative design and include impact measures that are beyond user satisfaction, self-reported and short-term based. Universities and educators should implement online assessment with careful planning and design more sophisticated ways that enhances and transforms assessment practices.

Abbreviations

RQ: research question; SAMR: substitution augmentation modification and redefinition model.

Acknowledgements

The authors would like to thank Dr Chandima Daskon for assistance in data extraction, and the two anonymous reviewers for their comments.

Authors' contributions

QL: Conceptualisation; Methodology; Formal analysis; Investigation; Resources; Data curation; Writing – Original draft; Writing – Review & editing; Visualisation; Supervision; Project administration. AH: Validation; Investigation; Writing – Review & editing. BD: Conceptualisation; Writing- Review & editing.

Authors' information

Qian Liu, ¹ Higher Education Development Centre, University of Otago, Dunedin, New Zealand, 9016.
Email: qian.liu@otago.ac.nz; ORCID: 0000-0002-1412-0615.

Anjin Hu, ¹ Higher Education Development Centre, University of Otago, Dunedin, New Zealand, 9016; ² School of International Studies, Hunan Institute of Technology, Hunan, China, 421002. Email: anjin.hu@postgrad.otago.ac.nz; ORCID: 0000-0002-3810-920X.

Ben Daniel, ¹ Higher Education Development Centre, University of Otago, Dunedin, New Zealand, 9016.
Email: ben.daniel@otago.ac.nz; ORCID: 0000-0002-1173-8225.

Funding

Not applicable.

Availability of data and materials

Data extraction results are openly available at OSF from
<https://osf.io/gzsqb/files/osfstorage/654c1fe8c9ac3f07302f7219>

Declarations**Competing interests**

The authors declare that they have no competing interests.

Author details

¹ Higher Education Development Centre, University of Otago, New Zealand

² School of International Studies, Hunan Institute of Technology, China

Received: 9 November 2023 Accepted: 17 April 2024

Published online: 1 January 2025 (Online First: 14 May 2024)

References

- Abdalla, R., Bishop, S. S., Villasante-Tezanos, A. G., & Bertoli, E. (2020). Comparison between students' self-assessment, and visual and digital assessment techniques in dental anatomy wax-up grading. *European Journal of Dental Education*, 25(3), 524–535. <https://doi.org/10.1111/eje.12628>
- Alden, J. (2011). Assessment of individual student performance in online team projects. *Journal of Asynchronous Learning Networks*, 15(3), 5–20. <https://doi.org/10.24059/olj.v15i3.193>
- Alessio, H. M., Malay, N., Maurer, K., Bailer, A. J., & Rubin, B. (2018). Interaction of proctoring and student major on online test performance. *International Review of Research in Open and Distributed Learning*, 19(5), 165–185. <https://doi.org/10.19173/irrodl.v19i5.3698>
- Alvarez, I., Espasa, A., & Guasch, T. (2012). The value of feedback in improving collaborative writing assignments in an online learning environment. *Studies in Higher Education*, 37(4), 387–400. <https://doi.org/10.1080/03075079.2010.510182>
- Avila, J., Sostmann, K., Breckwoldt, J., & Peters, H. (2016). Evaluation of the free, open source software WordPress as electronic portfolio system in undergraduate medical education. *BMC Medical Education*, 16, Article 157. <https://doi.org/10.1186/s12909-016-0678-1>
- Balta, N., Perera-Rodriguez, V. H., & Hervás-Gómez, C. (2018). Using socrative as an online homework platform to increase students' exam scores. *Education and Information Technologies*, 23(2), 837–850. <https://doi.org/10.1007/s10639-017-9638-6>
- Balter, O., Enstrom, E., & Klingenberg, B. (2013). The effect of short formative diagnostic web quizzes with minimal feedback. *Computers & Education*, 60(1), 234–242. <https://doi.org/10.1016/j.compedu.2012.08.014>
- Bausili, A. V. (2018). From piloting e-submission to electronic management of assessment (EMA): Mapping grading journeys. *British Journal of Educational Technology*, 49(3), 463–478. <https://doi.org/10.1111/bjet.12547>
- Bayrak, F. (2021). Investigation of the web-based self-assessment system based on assessment analytics in terms of perceived self-intervention. *Technology Knowledge and Learning*, 27(2022), 639–662. <https://doi.org/10.1007/s10758-021-09511-8>
- Bearman, M., Nieminen, J. H., & Ajajawi, R. (2023). Designing assessment in a digital world: An organising framework. *Assessment & Evaluation in Higher Education*, 48(3), 291–304. <https://doi.org/10.1080/02602938.2022.2069674>

- Beckman, K., Apps, T., Bennett, S., Dalgarno, B., Kennedy, G., & Lockyer, L. (2021). Self-regulation in open-ended online assignment tasks: The importance of initial task interpretation and goal setting. *Studies in Higher Education*, 46(4), 821–835. <https://doi.org/10.1080/03075079.2019.1654450>
- Biggs, J. (1996). Enhancing teaching through constructive alignment. *Higher Education*, 32(3), 347–364. <https://doi.org/10.1007/BF00138871>
- Binder, R., Friedli, A., & Fuentes-Afflick, E. (2016). The new academic environment and faculty misconduct. *Academic Medicine*, 91(2), 175–179. <https://doi.org/10.1097/acm.0000000000000956>
- Bleasel, J., Burgess, A., Weeks, R., & Haq, I. (2016). Feedback using an ePortfolio for medicine long cases: Quality not quantity. *BMC Medical Education*, 16, Article 278. <https://doi.org/10.1186/s12909-016-0801-3>
- Boitshwarelo, B., Reedy, A., & Billany, T. (2017). Envisioning the use of online tests in assessing twenty-first century learning: A literature review. *Research & Practice in Technology Enhanced Learning*, 12, 16. <https://doi.org/10.1186/s41039-017-0055-7>
- Boud, D., Dawson, P., Bearman, M., Bennett, S., Joughin, G., & Molloy, E. (2018). Reframing assessment research: Through a practice perspective. *Studies in Higher Education*, 43(7), 1107–1118. <https://doi.org/10.1080/03075079.2016.1202913>
- Brink, R., & Lautenbach, G. (2011). Electronic assessment in higher education. *Educational Studies*, 37(5), 503–512. <https://doi.org/10.1080/03055698.2010.539733>
- Buckley, E., & Cowap, L. (2013). An evaluation of the use of Turnitin for electronic submission and marking and as a formative feedback tool from an educator's perspective. *British Journal of Educational Technology*, 44(4), 562–570. <https://doi.org/10.1111/bjet.12054>
- Butler-Henderson, K., & Crawford, J. (2020). A systematic review of online examinations: A pedagogical innovation for scalable authentication and integrity. *Computers & Education*, 159, 104024. <https://doi.org/10.1016/j.compedu.2020.104024>
- Calderon, O., & Sood, C. (2020). Evaluating learning outcomes of an asynchronous online discussion assignment: A post-priori content analysis. *Interactive Learning Environments*, 28(1), 3–17. <https://doi.org/10.1080/10494820.2018.1510421>
- Carr, M., Prendergast, M., Breen, C., & Faulkner, F. (2017). How well do engineering students retain core mathematical knowledge after a series of high threshold online mathematics tests? *Teaching Mathematics and Its Applications*, 36(3), 136–150. <https://doi.org/10.1093/teamat/hrw021>
- Cerutti, B., Stollar, F., Escher, M., Blondon, K., Aujesky, S., Nendaz, M., & Galetto-Lacour, A. (2019). Comparison of long-menu and single-best-answer multiple choice questions in computer-based summative assessments: A randomised controlled trial. *BMC Medical Education*, 19, Article 219. <https://doi.org/10.1186/s12909-019-1651-6>
- Chen, H. Y., & Chuang, C. H. (2012). The learning effectiveness of nursing students using online testing as an assistant tool: A cluster randomized controlled trial. *Nurse Education Today*, 32(3), 208–213. <https://doi.org/10.1016/j.nedt.2011.03.004>
- Cheng, K. H., & Hou, H. T. (2015). Exploring students' behavioural patterns during online peer assessment from the affective, cognitive, and metacognitive perspectives: A progressive sequential analysis. *Technology Pedagogy and Education*, 24(2), 171–188. <https://doi.org/10.1080/1475939x.2013.822416>
- Cheng, K. H., & Tsai, C. C. (2012). Students' interpersonal perspectives on, conceptions of and approaches to learning in online peer assessment. *Australasian Journal of Educational Technology*, 28(4), 599–618. <https://doi.org/10.14742/ajet.830>
- Cheung, Y. L. (2016). A comparative study of paper-and-pen versus computer-delivered assessment modes on students' writing quality: A Singapore study. *Asia-Pacific Education Researcher*, 25(1), 23–33. <https://doi.org/10.1007/s40299-015-0229-2>
- Chew, E., Snee, H., & Price, T. (2016). Enhancing international postgraduates' learning experience with online peer assessment and feedback innovation. *Innovations in Education and Teaching International*, 53(3), 247–259. <https://doi.org/10.1080/14703297.2014.937729>
- Christensen Hughes, J. M., & McCabe, D. L. (2006). Understanding academic misconduct. *The Canadian Journal of Higher Education*, 36(1), 49–63.
- Conole, G., & Warburton, B. (2005). A Review of computer-assisted assessment. *Research in Learning Technology*, 13(1), 17–31. <https://doi.org/10.1080/0968776042000339772>
- Craft, J., & Ainscough, L. (2015). Development of an electronic role-play assessment initiative in bioscience for nursing students. *Innovations in Education and Teaching International*, 52(2), 172–184. <https://doi.org/10.1080/14703297.2014.931241>
- Crompton, H., & Burke, D. (2020). Mobile learning and pedagogical opportunities: A configurative systematic review of PreK-12 research using the SAMR framework. *Computers & Education*, 156, 103945. <https://doi.org/https://doi.org/10.1016/j.compedu.2020.103945>
- Daffin, L. W., & Jones, A. A. (2018). Comparing student performance on proctored and non-proctored exams in online psychology courses. *Online Learning*, 22(1), 131–145. <https://doi.org/10.24059/oli.v22i1.1079>
- Daniels, L. M., & Gierl, M. J. (2017). The impact of immediate test score reporting on university students' achievement emotions in the context of computer-based multiple-choice exams. *Learning and Instruction*, 52, 27–35. <https://doi.org/10.1016/j.learninstruc.2017.04.001>
- Darr, C. (2023). *Reliability and validity*. Ministry of Education, New Zealand Government. <https://assessment.tki.org.nz/Using-evidence-for-learning/Working-with-data/Concepts/Reliability-and-validity>

- Davies, A., Pantzopoulos, K., & Gray, K. (2011). Emphasising assessment 'as' learning by assessing wiki writing assignments collaboratively and publicly online. *Australasian Journal of Educational Technology*, 27(5), 798–812. <https://doi.org/10.14742/ajet.932>
- Deutsch, T., Herrmann, K., Frese, T., & Sandholzer, H. (2012). Implementing computer-based assessment - A web-based mock examination changes attitudes. *Computers & Education*, 58(4), 1068–1075. <https://doi.org/10.1016/j.compedu.2011.11.013>
- Ebadi, S., & Rahimi, M. (2019). Mediating EFL learners' academic writing skills in online dynamic assessment using Google Docs. *Computer Assisted Language Learning*, 32(5–6), 527–555. <https://doi.org/10.1080/09588221.2018.1527362>
- Ebadi, S., Weisi, H., Monkarezi, H., & Bahramlou, K. (2018). Exploring lexical inferencing as a vocabulary acquisition strategy through computerized dynamic assessment and static assessment. *Computer Assisted Language Learning*, 31(7), 790–817. <https://doi.org/10.1080/09588221.2018.1451344>
- Fang, J. W., Chang, S. C., Hwang, G. J., & Yang, G. (2021). An online collaborative peer-assessment approach to strengthening pre-service teachers' digital content development competence and higher-order thinking tendency. *Educational Technology Research and Development*, 69(2), 1155–1181. <https://doi.org/10.1007/s11423-021-09990-7>
- Fernando, W. (2018). Show me your true colours: Scaffolding formative academic literacy assessment through an online learning platform. *Assessing Writing*, 36, 63–76. <https://doi.org/10.1016/j.asw.2018.03.005>
- Fiddes, D. J., Korabinski, A. A., McGuire, G. R., Youngson, M. A., & McMillan, D. (2002). Are mathematics exam results affected by the mode of delivery. *Research in Learning Technology*, 10(6), 1–9.
- Figuerola-Canas, J., & Sancho-Vinuesa, T. (2021). Investigating the relationship between optional quizzes and final exam performance in a fully asynchronous online calculus module. *Interactive Learning Environments*, 29(1), 33–43. <https://doi.org/10.1080/10494820.2018.1559864>
- Flavin, M. (2021). A Disruptive Innovation perspective on students' opinions of online assessment. *Research in Learning Technology*, 29. <https://doi.org/10.25304/rlt.v29.2611>
- Formanek, M., Wenger, M. C., Buxner, S. R., Impey, C. D., & Sonam, T. (2017). Insights about large-scale online peer assessment from an analysis of an astronomy MOOC. *Computers & Education*, 113, 243–262. <https://doi.org/10.1016/j.compedu.2017.05.019>
- Geertshuis, S., Wass, R., & Liu, Q. (2022). Equipping graduates with future-ready capabilities: An application of learning theories to higher education. *Teaching in Higher Education*, 1–20. <https://doi.org/10.1080/13562517.2022.2137399>
- Gikandi, J. W., Morrow, D., & Davis, N. E. (2011). Online formative assessment in higher education: A review of the literature. *Computers & Education*, 57(4), 2333–2351. <https://doi.org/10.1016/j.compedu.2011.06.004>
- Grainger, R., Liu, Q., & Gladman, T. (2024). Learning technology in health professions education: Realising an (un) imagined future. *Medical Education*, 58(1), 36–46. <https://doi.org/10.1111/medu.15185>
- Grant, M. J., & Booth, A. (2009). A typology of reviews: An analysis of 14 review types and associated methodologies. *Health Information & Libraries Journal*, 26(2), 91–108. <https://doi.org/10.1111/j.1471-1842.2009.00848.x>
- Gray, K., Waycott, J., Clerehan, R., Hamilton, M., Richardson, J., Sheard, J., & Thompson, C. (2012). Worth it? Findings from a study of how academics assess students' Web 2.0 activities. *Research in Learning Technology*, 20(1). <https://doi.org/10.3402/rlt.v20i0.16153>
- Grieve, R., Padgett, C. R., & Moffitt, R. L. (2016). Assignments 2.0: The role of social presence and computer attitudes in student preferences for online versus offline marking. *Internet and Higher Education*, 28, 8–16. <https://doi.org/10.1016/j.iheduc.2015.08.002>
- Haack, S., Fornoff, A., Caligiuri, F., Dy-Boarman, E., Bottenberg, M., Mobley-Bukstein, W., Bryant, G., & Bryant, A. (2017). Comparison of electronic versus paper rubrics to assess patient counseling experiences in a skills-based lab course. *Currents in Pharmacy Teaching and Learning*, 9(6), 1117–1122. <https://doi.org/10.1016/j.cptl.2017.07.001>
- Hamilton, E. R., Rosenberg, J. M., & Akcaoglu, M. (2016). The substitution augmentation modification redefinition (SAMR) model: A critical review and suggestions for its use. *TechTrends*, 60(5), 433–441. <https://doi.org/10.1007/s11528-016-0091-y>
- Hay, P. J., Engstrom, C., Green, A., Friis, P., Dickens, S., & Macdonald, D. (2013). Promoting assessment efficacy through an integrated system for online clinical assessment of practical skills. *Assessment & Evaluation in Higher Education*, 38(5), 520–535. <https://doi.org/10.1080/02602938.2012.658019>
- Hensiek, S., DeKorver, B. K., Harwood, C. J., Fish, J., O'Shea, K., & Towns, M. (2016). Improving and assessing student hands-on laboratory skills through digital badging. *Journal of Chemical Education*, 93(11), 1847–1854. <https://doi.org/10.1021/acs.jchemed.6b00234>
- Herbert, K., Demskoi, D., & Cullis, K. (2019). Creating mathematics formative assessments using LaTeX, PDF forms and computer algebra. *Australasian Journal of Educational Technology*, 35(5), 153–167. <https://doi.org/10.14742/ajet.4539>
- Ho, V. W., Harris, P. G., Kumar, R. K., & Velan, G. M. (2018). Knowledge maps: A tool for online assessment with automated feedback. *Medical Education Online*, 23, Article 1457394. <https://doi.org/10.1080/10872981.2018.1457394>

- Hong, Q. N., Gonzalez-Reyes, A., & Pluye, P. (2018). Improving the usefulness of a tool for appraising the quality of qualitative, quantitative and mixed methods studies, the Mixed Methods Appraisal Tool (MMAT). *Journal of Evaluation in Clinical Practice*, 24(3), 459–467. <https://doi.org/10.1111/jep.12884>
- Huwendiek, S., Reichert, F., Duncker, C., de Leng, B. A., van der Vleuten, C. P. M., Muijtjens, A. M. M., Bosse, H.-M., Haag, M., Hoffmann, G. F., Tönshoff, B., & Dolmans, D. (2017). Electronic assessment of clinical reasoning in clerkships: A mixed-methods comparison of long-menu key-feature problems with context-rich single best answer questions. *Medical Teacher*, 39(5), 476–485. <https://doi.org/10.1080/0142159X.2017.1297525>
- Hwang, W. Y., Hsu, J. L., Shadieff, R., Chang, C. L., & Huang, Y. M. (2015). Employing self-assessment, journaling, and peer sharing to enhance learning from an online course. *Journal of Computing in Higher Education*, 27(2), 114–133. <https://doi.org/10.1007/s12528-015-9096-3>
- Kaliyadan, F., Khan, A., Kuruvilla, J., & Feroze, K. (2014). Validation of a computer based objective structured clinical examination in the assessment of undergraduate dermatology courses. *Indian Journal of Dermatology, Venereology and Leprology*, 80(2), 134–136. <https://doi.org/10.4103/0378-6323.129386>
- Kamrood, A. M., Davoudi, M., Ghaniabadi, S., & Amirian, S. M. R. (2019). Diagnosing L2 learners' development through online computerized dynamic assessment. *Computer Assisted Language Learning*, 34(7), 868–897. <https://doi.org/10.1080/09588221.2019.1645181>
- Khalaf, K., El-Kishawi, M., Moufti, M. A., & Al Kawas, S. (2020). Introducing a comprehensive high-stake online exam to final-year dental students during the COVID-19 pandemic and evaluation of its effectiveness. *Medical Education Online*, 25(1), Article 1826861. <https://doi.org/10.1080/10872981.2020.1826861>
- Krathwohl, D. R. (2002). A revision of Bloom's taxonomy: An overview. *Theory into Practice*, 41(4), 212–218. https://doi.org/10.1207/s15430421tip4104_2
- Lai, C. Y., Chen, L. J., Yen, Y. C., & Lin, K. Y. (2020). Impact of video annotation on undergraduate nursing students' communication performance and commenting behaviour during an online peer-assessment activity. *Australasian Journal of Educational Technology*, 36(2), 71–88. <https://doi.org/10.14742/ajet.4341>
- Lee, K., & Fanguy, M. (2022). Online exam proctoring technologies: Educational innovation or deterioration? *British Journal of Educational Technology*, 53(3), 475–490. <https://doi.org/10.1111/bjet.13182>
- Levy, Y., Ramim, M. M., Furnell, S. M., & Clarke, N. L. (2011). Comparing intentions to use university-provided vs vendor-provided multibiometric authentication in online exams. *Campus - Wide Information Systems*, 28(2), 102–113. <https://doi.org/10.1108/10650741111117806>
- Lim, A. S., Lee, S. W. H., Karunaratne, N., & Caliph, S. (2020). Pharmacy students' perceptions and performance on the use of an online virtual experience tool for practicing objective structured clinical examinations. *American Journal of Pharmaceutical Education*, 84(11), 1467–1475, Article 7920. <https://doi.org/10.5688/ajpe7920>
- Lin, C. J. (2019). An online peer assessment approach to supporting mind-mapping flipped learning activities for college English writing courses. *Journal of Computers in Education*, 6(3), 385–415. <https://doi.org/10.1007/s40692-019-00144-6>
- Lin, J. W., & Lai, Y. C. (2013). Online formative assessments with social network awareness. *Computers & Education*, 66, 40–53. <https://doi.org/10.1016/j.compedu.2013.02.008>
- Liu, I. F., Chen, R. S., & Lu, H. C. (2015). An exploration into improving examinees' acceptance of participation in an online exam. *Educational Technology & Society*, 18(2), 153–165.
- Liu, Q., Wald, N., Daskon, C., & Harland, T. (2023). Multiple-choice questions (MCQs) for higher-order cognition: Perspectives of university teachers. *Innovations in Education and Teaching International*, 1–13. <https://doi.org/10.1080/14703297.2023.2222715>
- Marinho, P., Fernandes, P., & Pimentel, F. (2021). The digital portfolio as an assessment strategy for learning in higher education. *Distance Education*, 42(2), 253–267. <https://doi.org/10.1080/01587919.2021.1911628>
- Mina, E. E. S., Eifert, C., Ireland, M., Fine, C., Micevski, V., Wojtiuk, R., Valderrama, M., & Wilson, G. (2011). The development of an online instrument for prior learning assessment and recognition of internationally educated nurses: A pilot study. *International Review of Research in Open and Distance Learning*, 12(1), 100–110. <https://doi.org/10.19173/irrodl.v12i1.896>
- Mostert, M., & Snowball, J. D. (2013). Where angels fear to tread: Online peer-assessment in a large first-year class. *Assessment & Evaluation in Higher Education*, 38(6), 674–686. <https://doi.org/10.1080/02602938.2012.683770>
- Nehm, R. H., & Haertig, H. (2012). Human vs. computer diagnosis of students' natural selection knowledge: Testing the efficacy of text analytic software. *Journal of Science Education and Technology*, 21(1), 56–73. <https://doi.org/10.1007/s10956-011-9282-7>
- Nielsen, W., Georgiou, H., Jones, P., & Turney, A. (2020). Digital explanation as assessment in university science. *Research in Science Education*, 50(6), 2391–2418. <https://doi.org/10.1007/s11165-018-9785-9>
- Noroozi, O., Biemans, H., & Mulder, M. (2016). Relations between scripted online peer feedback processes and quality of written argumentative essay. *Internet and Higher Education*, 31, 20–31. <https://doi.org/10.1016/j.iheduc.2016.05.002>
- Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., & Brennan, S. E. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ*, 372, 71. <https://doi.org/10.1136/bmj.n71>
- Paredes, S. G., Pena, F. D. J., & Alcazar, J. M. D. (2021). Remote proctored exams: Integrity assurance in online education? *Distance Education*, 42(2), 200–218. <https://doi.org/10.1080/01587919.2021.1910495>

- Pezzino, M. (2018). Online assessment, adaptive feedback and the importance of visual learning for students. The advantages, with a few caveats, of using MapleTA. *International Review of Economics Education*, 28, 11–28. <https://doi.org/10.1016/j.iree.2018.03.002>
- Popay, J., Roberts, H., Sowden, A., Petticrew, M., Arai, L., Rodgers, M., Britten, N., Roen, K., & Duffy, S. (2006). Guidance on the conduct of narrative synthesis in systematic reviews. *A product from the ESRC methods programme Version, 1*, b92. <https://doi.org/10.13140/2.1.1018.4643>
- Reilly, E. D., Stafford, R. E., Williams, K. M., & Corliss, S. B. (2014). Evaluating the validity and applicability of automated essay scoring in two massive open online courses. *International Review of Research in Open and Distributed Learning*, 15(5), 83–98. <https://doi.org/10.19173/irrodl.v15i5.1857>
- Rivera-Mata, J. (2021). How to Teach Online? Recommendations for the assessment of online exams with University students in the USA in times of pandemic. *International Journal of Educational Research and Innovation*, 15, 188–202. <https://doi.org/10.46661/ijeri.5003>
- Ros, N. F., Lucena, F., Inarrairaegui, M., Landecho, M. F., Sunsundegui, P., Jordan-Iborra, C., Pineda, I., Quiroga, J., & Herrero, J. I. (2021). Web-based formative assessment through clinical cases: Role in pathophysiology teaching. *BMC Medical Education*, 21(1), Article 249. <https://doi.org/10.1186/s12909-021-02691-y>
- Sebok-Syer, S. S., Goldszmidt, M., Watling, C. J., Chahine, S., Venance, S. L., & Lingard, L. (2019). Using electronic health record data to assess residents' clinical performance in the workplace: The good, the bad, and the unthinkable. *Academic Medicine*, 94(6), 853–860. <https://doi.org/10.1097/acm.0000000000002672>
- Snodgrass, S. J., Ashby, S. E., Rivett, D. A., & Russell, T. (2014). Implementation of an electronic Objective Structured Clinical Exam for assessing practical skills in pre-professional physiotherapy and occupational therapy programs: Examiner and course coordinator perspectives. *Australasian Journal of Educational Technology*, 30(2), 152–166. <https://doi.org/10.14742/ajet.348>
- St-Onge, C., Ouellet, K., Lakhali, S., Dubé, T., & Marceau, M. (2022). COVID-19 as the tipping point for integrating e-assessment in higher education practices. *British Journal of Educational Technology*, 53(2), 349–366. <https://doi.org/10.1111/bjet.13169>
- Swart, A. J., & Meda, L. (2021). Online self-assessments need to be regular and reflective to foster student engagement and success. *International Journal of Electrical Engineering Education*, Article 0020720921997674. <https://doi.org/10.1177/0020720921997674>
- Thomas, J. A., Wadsworth, D., Jin, Y., Clarke, J., Page, R., & Thunders, M. (2017). Engagement with online self-tests as a predictor of student success. *Higher Education Research & Development*, 36(5), 1061–1071. <https://doi.org/10.1080/07294360.2016.1263827>
- Thompson, M. M., & Braude, E. J. (2016). Evaluation of Knowla: An online assessment and learning tool. *Journal of Educational Computing Research*, 54(4), 483–512. <https://doi.org/10.1177/0735633115621923>
- Tinoca, L., & Oliveira, I. (2013). Formative assessment of teachers in the context of an online learning environment. *Teachers and Teaching*, 19(2), 214–227. <https://doi.org/10.1080/13540602.2013.741836>
- Wang, Y. (2019). Why students have conflicts in peer assessment? An empirical study of an online peer assessment community. *Sustainability*, 11(23), 6807. <https://doi.org/http://dx.doi.org/10.3390/su11236807>
- Way, K. A., Burrell, L., D'Allura, L., & Ashford-Rowe, K. (2021). Empirical investigation of authentic assessment theory: An application in online courses using mimetic simulation created in university learning management ecosystems. *Assessment & Evaluation in Higher Education*, 46(1), 17–35. <https://doi.org/10.1080/02602938.2020.1740647>
- Whitworth, D. E., & Wright, K. (2015). Online assessment of learning and engagement in university laboratory practicals. *British Journal of Educational Technology*, 46(6), 1201–1213. <https://doi.org/10.1111/bjet.12193>
- William, D. (2011). What is assessment for learning? *Studies in Educational Evaluation*, 37(1), 3–14. <https://doi.org/10.1016/j.stueduc.2011.03.001>
- Wilson, K., Boyd, C., Chen, L. W., & Jamal, S. (2011). Improving student performance in a first-year geography course: Examining the importance of computer-assisted formative assessment. *Computers & Education*, 57(2), 1493–1500. <https://doi.org/10.1016/j.compedu.2011.02.011>
- Yang, Y., & Qian, D. D. (2017). Assessing English reading comprehension by Chinese EFL learners in computerized dynamic assessment. *Language Testing in Asia*, 7(1), 1–15. <https://doi.org/10.1186/s40468-017-0042-3>
- Yasuda, J., Mae, N., Hull, M. M., & Taniguchi, M. (2021). Optimizing the length of computerized adaptive testing for the Force Concept Inventory. *Physical Review Physics Education Research*, 17(1), Article 010115. <https://doi.org/10.1103/PhysRevPhysEducRes.17.010115>
- Yu, F. Y., & Wu, C. P. (2016). The effects of an online student-constructed test strategy on knowledge construction. *Computers & Education*, 94, 89–101. <https://doi.org/10.1016/j.compedu.2015.11.005>
- Zhan, Y. (2021). What matters in design? Cultivating undergraduates' critical thinking through online peer assessment in a Confucian heritage context. *Assessment & Evaluation in Higher Education*, 46(4), 615–630. <https://doi.org/10.1080/02602938.2020.1804826>
- Zheng, L. Q., Cui, P. P., Li, X., & Huang, R. H. (2018). Synchronous discussion between assessors and assesseees in web-based peer assessment: Impact on writing performance, feedback quality, meta-cognitive awareness and self-efficacy. *Assessment & Evaluation in Higher Education*, 43(3), 500–514. <https://doi.org/10.1080/02602938.2017.1370533>
- Zlatovic, M., Balaban, I., & Kermek, D. (2015). Using online assessments to stimulate learning strategies and achievement of learning goals. *Computers & Education*, 91, 32–45. <https://doi.org/10.1016/j.compedu.2015.09.012>

Publisher's Note

The Asia-Pacific Society for Computers in Education (APSCE) remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Research and Practice in Technology Enhanced Learning (RPTEL)
is an open-access journal and free of publication fee.