JOURNAL OF EDUCATIONAL RESEARCH & PRACTICE

Journal of Educational Research & Practice
2023, Volume 13, Issue 1, Pages 476–496
DOI: 10.5590/JERAP.2023.13.1.31

Original Research WALDEN UNIVERSITY

Online Assessment in Large Undergraduate Courses During COVID-19 Emergency Response Teaching

Kate Maloney Williams, MSc
University of Maryland, College Park, Maryland, United States
https://orcid.org/0000-0002-3770-6678

Alice E. Donlan, PhD
University of Maryland, College Park, Maryland, United States
https://orcid.org/0000-0001-8776-422X

Contact: *kmwills@umd.edu*

Abstract

The transition to online instruction during the COVID-19 pandemic was unprecedented and forced many universities to quickly embrace online distance learning. This context created new challenges, particularly around assessment strategies. Empirical research has demonstrated that formative assessment fosters more active learning in online classrooms. However, formative assessment strategies are not always adapted well to online platforms based on the nature of the subject matter and the size of the class. This qualitative case study sought to understand instructors' experiences and strategies for conducting assessment remotely, specifically for large-size undergraduate courses. The investigation relied on data from semi-structured interviews with University of Maryland, College Park instructors who received a Teaching Innovation Grant from the Provost's Office in Summer 2020 intended to fund sustainable online delivery beyond the emergency response teaching phase. For this analysis, we analyzed the transcripts of 13 interviews, representing a diverse range of programs, schools, and faculty seniority levels at the university. Findings show instructors experienced several successes during course retooling, including significant increases in student performance. Most instructors also indicated that they would continue to keep new online assessment strategies for the future, regardless of whether that future includes online, blended, or in-person delivery. Despite the anticipation that the pandemic would fuel more opportunities for cheating, there was only one experience of academic dishonesty.

Keywords: *distance education and online learning, evaluation methodologies, pedagogical issues, teaching/learning strategies, 21st-century abilities*

Date Submitted: June 23, 2023 | **Date Published:** November 21, 2023

Recommended Citation

Williams, K. M., & Donlan, A. E. (2023). Online assessment in large undergraduate courses during COVID-19 emergency response teaching. *Journal of Educational Research and Practice, 13*, 476–496. https://doi.org/10.5590/JERAP.2023.13.1.31

Introduction

The transition to online instruction during the COVID-19 pandemic was unprecedented and stressful for college instructors. It forced many universities to quickly implement online distance learning due to social distancing mandates (Barros & Joung, 2020; Witze, 2020). This context created new challenges, such as selecting the most appropriate platforms and instructional methods to increase access to and the quality of remote delivery (Costa et al., 2021). At the beginning of the pandemic, most instructors focused on basic content transfer and emergency response teaching with little attention paid to new assessment strategies better suited for online learning (Eaton, 2020). However, as the pandemic continued, there was more need to create sustainable online learning environments and assess complex learning outcomes (Korkmaz & Toraman, 2020; Olsen, 2021; Rapanta et al., 2020). Instructors not only needed to adapt their assessment strategies for an online or blended course delivery; they also needed to adapt their assessment strategies to be more flexible and understanding of student social–emotional welfare during a traumatic global emergency (INEE, 2018; Kamei & Harriott, 2021; Katzman & Stanton, 2020; Tunc et al., 2022; Williams & Corwith, 2021; Zieher et al, 2021).

In Summer 2020, the University of Maryland, College Park used funding from the U.S. Coronavirus Aid, Relief, and Economic Security (CARES) Act to award hundreds of Innovation Grants to faculty across its schools, departments, and programs. The grants provided financial support for the time and materials required to retool coursework for more long-term remote delivery. Instructors used the funding for many purposes, including summer salaries to create new lecture videos, syllabi, digital spaces, and assessments, as well as purchase recording equipment and software licenses. One common concern from instructors on campus surrounded the need for rigorous assessments of learning that could be done from home without proctoring. The aim of this study is to document themes from instructors on their decision-making processes regarding assessment changes in their large-size university courses and determine approaches that may be impactful for future delivery and scale-up in any technology-enabled classroom. Our study contributes to the research base on promoting sustainable online learning environments and meaningful learning outcomes in emergencies, as well as overall quality improvement in higher education.

Literature Review

The best online courses allow students to interact with content through multiple modes of expression and assessment (CAST, 2018; Erlandson, 2002; Wyatt-Smith & Kimber, 2009). The design of online courses is particularly important because they require more cognitive, metacognitive, and social skills to navigate than in-person classes (Conrad & Openo, 2018; Webb & Gibson, 2015). This increased demand on the student requires a targeted focus on education goals and outcomes-based approaches during the course design. Since online classes demand these increased skills, the question arises: should the nature of online assessment stay the same as in-person learning? Assessment in online learning contexts can present more challenges than face-to-face contexts (Vonderwell et al., 2007). These include more misalignment with curriculum priorities, an over-reliance on automated grading, and negative impacts on student confidence and well-being, which are particularly vulnerable during education in emergencies (Attwood & Radnofsky, 2007; Deneen & Boud, 2014; Gee & Shaffer, 2010; Schwartz & Arena, 2009; Shute & Kim, 2014; Timmis et al., 2015).

Research has indicated that conventional methods of assessment (e.g., closed note, individual exams) are increasingly misaligned with the demands of the 21st century and for creating communities of shared practice (Boud & Molloy, 2013; Elwood & Klenowski, 2002; Hattie & Timperley, 2007; Nicol & McFarlane-Dick, 2006; Timmis et al., 2015). Most educational assessment still targets the types of skills that were seen as important throughout the 20th century—storing relevant information and retrieving it upon demand (Care & Vista, 2017). However, in modern workplaces these skills are largely supplanted by the computer and the Internet, media specifically designed for rapid access to information. In the 21st century, career readiness means equipping

students with a nuanced set of skills to prepare them for an unknown workforce, focused on generativity and innovation, collaboration and participation, computational thinking, communication and soft skills, and information and digital literacy (González-Pérez & Ramírez-Montoya, 2022). Higher education and business leaders cite soft skills as being more important drivers of success than traditional conceptions of literacy, numeracy, and rote learning (Buckle, n.d.). Moreover, contemporary assessment needs to offer complex and authentic activities that engage learners in decision-making and problem-solving relevant to real-world situations and metacognition (Crisp & Ward, 2008; Lin, 2008; Mackey, 2009; Wang et al., 2008). Digital tools provide affordances that allow for these constructivist modes of education and inquiry (Timmis et al., 2015), including the use of formative assessment and project-based learning as methods for summative evaluation.

For many who have never taught online before, the sudden pivot to emergency remote learning during COVID-19 provided "little time to consider how assessment in e-learning can, and should differ from how we assess students in face-to-face environments" (Eaton, 2020, p. 82). It is important to remember that what gets measured matters. Defining what one intends to measure in course assessments and establishing a consistent system to define the magnitude of what is being measured are crucial to assessment; these choices are often straightforward for rote learning but not for measures of higher-order thinking (Care & Vista, 2017).

Academic Dishonesty and Cheating

One major hurdle instructors face when designing online assessments is preventing cheating. Although, it is important to note that empirical evidence has found less academic misconduct in online courses compared to traditional delivery (Davis et al., 2009; Hart & Morgan, 2010; Oosterhof et al., 2008; Stuber-McEwen et al., 2009). Common strategies used to mediate online cheating include making tests "open-book, take-home" (Gamage et al., 2020). Additionally, these exams often contain one or more questions that require essay or long-form responses or the citation of specific examples and references. Other tactics might include creating multiple versions of exams, scrambling questions, and distributing exams in a pattern that ensures no students have the same version (Rajet & Stitzel, 2020). Other solutions include "modifications of exam structure (using different forms, one-way exams, reducing the number of questions/page[s], reduce time limits), chang[ing] the mode of assessment (oral exams, substituting exams with other forms of assessment)" (Elsalem et al., 2021, p. 330). Brouwer and McDonnell (2009) argue that projects, portfolios, self-assessments, peer evaluations, timed tests and quizzes, and asynchronous discussions were found to be effective assessments for online courses.

Lowering the Stakes of Each Assessment

Academic integrity is particularly at risk when students are under extreme stress, as they were during the COVID-19 pandemic and vulnerable to making poor choices regarding academic conduct (Eaton, 2020). Cheating is more likely to occur for high stakes assessments (Nichols & Berliner, 2007). Similarly, plagiarism can be fueled by the ease of technology and greater access to information (Baron & Crooks, 2005). Often, the recommended practice is for instructors to restructure their online course syllabi to be weighted more on participation, essay assignments, group work, and project-based learning. However, other strategies can include breaking exams into smaller, more frequent, lower stakes quizzes (Salas-Morera et al., 2012). This practice also falls under "concept chunking"—an effective strategy in online training courses for breaking up dense material into shorter, more manageable pieces to relieve screen fatigue and cognitive load (CAST, 2018). Existing evidence on Massive Open Online Classrooms (MOOCs) also promotes connectivist strategies, high instructor-student interaction, personalization, and an array of instructional practices and technology tools to support learning (Bartoletti, 2016; Bawa, 2016; Bonk et al., 2018; Hone & El Said, 2016; Zhu et al., 2018). Connectivism refers to theories that learning best occurs when peers freely exchange opinions, ideas, and other viewpoints through a collaborative process.

Additionally, formative assessment provides an alternative method to summative assessment than higher stakes exams. Formative assessment includes all activities that teachers conduct to provide feedback to students prior to a high-stakes assignment that allows students to practice content and teachers to adapt learning activities according to individual or collective performance (Sorensen & Takle, 2005). Its objective is to better understand how one's students receive and apply lesson material, focusing on the process of learning over the final product or outcomes (Ludwig-Hardman & Dunlap, 2003). It is also one of the leading strategies for online course assessment, particularly because it can be designed to negate cheating. Although it may rely on lower stakes or unweighted quizzes, a useful strategy is to rely on student reflections, whether oral or written. These reflection exercises allow teachers to generate personalized information that can be used to shape instruction and help students understand conceptual knowledge and advance successfully in their coursework (Black & Wiliam, 1998; Broadfoot et al., 2001; Gikandi et al., 2011; Shepard, 2000; Timmis et al., 2015). Empirical research demonstrates that formative assessment of this nature fosters more active learning with students and that technology inherently facilitates new opportunities for effective individualized data gathering (Looney, 2011). However, in order to be valid, feedback needs to be timely, ongoing, useful, and easy to understand (Gaytan & McEwen, 2007; Koh, 2008; Wang et al., 2008; Wolsey, 2008).

Utilizing Remote Online Proctoring Software

Another challenge that instructors face when conducting traditional exam-based assessments online is how to oversee strict proctoring or invigilation (Gamage et al., 2020). There are several ways that algorithmic assessment and Artificial Intelligence (AI) can drive strategies for online assessment. For one, Computerized Adaptive Testing (CAT) has become increasingly common in educational assessment in the United States, particularly for large-scale testing programs like the Graduate Record Examination (GRE) or MOOCs (Jian-quan et al., 2007; Meyer & Zhu, 2013). Many developers recognized the CAT procedure as a simple process: "A test-taker is estimated (or guessed) to have a certain ability. An item of the equivalent level of difficulty is asked. If the test-taker succeeds on the item, the ability estimate is raised. If the test-taker fails in the item, the ability estimate is lowered. Another item is asked, targeted on the revised ability estimate. And the process repeats" (Jian-quan et al., 2007, p. 77). In the online learning context, this readily allows personalized learning and both formative and summative evaluation (Guzmán & Conejo, 2005; Oppl et al., 2017; Salcedo et al., 2005).

Additionally, instructors have access to proctoring software such as Turnitin, IntegriGuard, EduTie, PlagiServ, which use artificial intelligence and algorithmic logic to identify cheating (Heberling, 2002). However, software also harbors limitations, including inaccuracies and the failure to account for what constitutes misconduct and plagiarism across various cultures and disciplines. According to Gamage et al. (2020), the types of practical assessments successfully moved online during COVID-19 include laboratory-based practicals (sciences), performance-based assessments (fine arts), physical artifact development (engineering), psychomotor skills (nursing), interpersonal skills, and language skills. These assessments relied on "video-based uploads using Cloud technology, online simulation tasks, submitting online portfolio, real-time observed practices/vivas via Zoom or Blackboard Collaborate" (Gamage et al., 2020, n.p.).

As an alternative to AI proctoring, some instructors also report requiring students to complete assignments and/or exams in front of a web camera, providing the instructor with live streaming images (Baron & Crooks, 2005; Carnevale, 1999). Other strategies can include curriculum rotation (Baron & Crooks, 2005) or the eight essential Online Exam Control Procedures (OECPs) developed by Cluskey et al. (2011).

Theoretical Framework

This paper is grounded in the SAMR (Substitution, Augmentation, Modification, Redefinition) Model (Puentedura, 2006), a framework that guides more transformative technology integration in classrooms. This perspective categorizes integrations of technology in teaching into four types: Substitution, Augmentation, Modification, and Redefinition. Substitution is exemplified by changes from teaching without technology to

teaching with technology with no substantive change in the functionality or benefit of the practice (e.g., moving from live polling with students raising their hands to polling software). The change becomes Augmentation when there is some functional improvement with the new tool (e.g., using automated graphs in the polling software to summarize students' responses). Next, the change becomes a Modification when the technology significantly redesigns the task (e.g., gathering poll data from multiple student questions and finding relationships among students' responses in live time). Finally, the instructor Redefines the learning activity with technology when it is used to create novel tasks that were not previously possible (e.g., sorting students by their poll responses into small discussion groups so that students can engage in a robust, recorded debate that the instructor could assess as an assignment in the class). The SAMR Model allows for critical reflection in lesson planning in relation to the purposeful use of technologies for quality learning—to whichever end, not excluding any form of assessment and evaluation. Ideally, every mediated classroom activity is held to this lens. However, it is currently unknown how most online emergency response teaching activities compare to this guidance and how the SAMR Model can help drive ongoing improvements to the resulting educational delivery.

Purpose of the Study, Research Questions, and Hypotheses

We are interested in applying the SAMR Model to the COVID-19 context because many instructors worldwide rapidly replaced face-to-face instruction with remote learning. In doing so, with instructors having minimal preparation or, in many instances, expertise, the quality and outcomes varied. Knowing where the most successful outcomes fell on the SAMR Model may also contribute to the evidence base on its validity. This understanding may better prepare educators to more proactively and sustainably design technology-enhanced assessment strategies for their courses. Though the COVID-19 pandemic was a surprise disruption, some instructors have used the necessary course redesign to include new, exploratory practices that benefit learning. However, we, the researchers, expect there to be heterogeneity in the pace of this transition and predict that the ways instructors revise their assessments for remote instruction will also vary in how they integrate technology and fall on the SAMR Model spectrum.

This investigation relied on data from semi-structured interviews with University of Maryland, College Park instructors who received a Teaching Innovation Grant from the Provost's Office in Summer 2020. Our primary objective was to better understand how instructors of large undergraduate classes during COVID-19 emergency remote teaching included novel assessment strategies compared to best practices in online teaching so that they could be better understood and replicated. Secondary objectives included illuminating instructors' decision-making processes with regard to assessment strategies in higher education and understanding where successful assessment strategies fell in relation to the SAMR Model of technology integration. Central research questions focused on:

1. What technology-enhanced assessment strategies did instructors implement in their large undergraduate classes during the Fall 2020 semester of the COVID-19 pandemic?
2. In what ways do these assessment strategies relate to the SAMR Model?
3. How do instructors reflect on the benefits and challenges of their assessment strategies and their plans for assessment in the future?

Methods

Participants and Context

An invitation to participate was emailed individually to instructors who had formerly participated in similarly themed surveys administered by the University of Maryland Teaching and Learning Transformation Center.

Interested individuals were then selected based on whether they had received an Innovation Grant in Summer 2020. We presumed that these individuals would be more likely to have had the leverage and support to try more innovative practices in their delivery based on this funding. Participants were also required to be over the age of 18 and actively teaching a course in Fall 2020. This yielded 41 volunteer participants, including graduate student instructors, professional track, tenure track, and tenured professors. We selected the 13 interviews to review for this analysis because they all spoke to the assessment experiences of instructors who were teaching large undergraduate classes, ranging from 30 to 340 students. We wanted to specifically target instructors' experiences and strategies for conducting assessments remotely, particularly for large-size undergraduate courses. The other 28 interviews conducted were excluded from the analysis because they were either related to small classes (less than 30 students) or graduate-level classes. For transparency of the diversity of our sample, we have included a breakdown of the instructors' demographics in Table 1.

Table 1. *Breakdown of Instructor Demographics (Assessment Sample)*

Gender	Race	Rank	Mean Age
9 Men	10 White	8 Professors	50–59 years old
4 Women	2 Asian or Asian American	5 Lecturers	
0 Nonbinary	1 Hispanic or Latinx		

Instrumentation and Data Collection

Due to the pandemic, all communications and interviews were virtual, completed and recorded over Zoom. Interviews were limited to 30 minutes in length. To facilitate conversations, we split interview questions into two categories: high and low priority. If conversations were running too long, low-priority questions were skipped to keep time. Time cuts occurred in almost all interviews. Occurrence of which low-priority questions were asked varied by the content of the conversation and the interviewer (See Table 2). All interview protocols were approved by the University of Maryland Institutional Review Board.

Table 2. *Matrix of Relevant Interview/Survey Questions*

General Experiences	Innovations	In Practice
What courses are you teaching this semester?	**How did you adapt your course during the COVID-19 pandemic?**	Have you felt that you've been able to connect with your students?
What course did you receive grant funding for?	**What did the funding allow you to do?**	Did you change the content of the course in any way?
What was your experience with this course before the pandemic? Have you taught it before? Did you ever teach it online?	Did you have support/wish you had support for this work?	Are you using synchronous instruction, asynchronous instruction, or both types of instruction?
What was your goal with the funding?	**What has gone well?**	**Has your assessment strategy changed at all?**
	What challenges remain? Did anything fail that you'd want to warn people about?	**What are you most proud of or excited about?**

Note: **Bold** denotes which questions were marked as a higher priority.

In order to protect confidentiality, all video recordings were destroyed and only the audio recordings were retained. However, interviewers also took notes during the interviews, which sometimes captured further details on body language and other nuances. These files and transcripts were stored on Box, a secure web-based data storage platform, in folders that only the principal investigator and co-investigators could access. In a demographics poll administered during the interviews, faculty were asked to indicate their college, gender, race, age, and faculty level but not their names or departments. The university intends to store de-identified data sets indefinitely after study completion.

Data Analysis

Qualitative researchers sometimes struggle to make an effective argument for their interpretation of data. They must tell a story rather than point to more binary indicators, which means that interpretation can ultimately appear subjective to the audience (Wolcott, 1994). To minimize this threat, data analysis relied heavily on three waves of thematic coding. The initial coding process did not use a coding framework but followed pattern coding. This relied on both the in vivo method and a "versus" approach with the latter seeking to identify in binary terms individuals, groups, institutions, phenomena, and processes that are seemingly in direct contrast to one another (Saldaña, 2016). As an inductive method, in vivo coding involves the organic coding of data to construct descriptive, higher-order themes. Here, we followed a three-stage process, which began with reading interview transcripts and other documents for phrases that seemed to be significant and noteworthy. For Stage 2, codes were consolidated and grouped according to differences and possible content overlap. For Stage 3, if the categories began to demonstrate both significant nuances as well as enough sources, they were then split into new categories to make those differences more readily identifiable (i.e., "challenges" were split into new codes, such as "software issues" or "student engagement"). For the purposes of this analysis, the sample size was then filtered down to only the 13 sampled interviews.

During the second and third stages of analysis, structural coding was used to review these 13 interviews. As a method, structural coding seeks to apply a content-based phrase or label representing a topic of inquiry to a segment of data that relates to a specific research question used to frame the interview (MacQueen et al., 2008; Saldaña, 2016). Further, experts suggest that a "code frequency report can help identify which themes, ideas, or domains were common and which rarely occurred" (Namey et al., 2008, p. 143), which was generated here for analysis.

Throughout research design, data collection, and reflection, we also created an audit trail, using checklists and annotations to analyze interview transcripts. These measures facilitated the coding process and use of NVivo software. NVivo was chosen over other qualitative analysis software due to the researchers' familiarity and licensing. However, as a tool, NVivo readily facilitates the analysis of unstructured interview data through thematic coding affordances, helping users note patterns, link connections, and reorganize structure for flexibility and rigor. In this way, it facilitates trend identification and cross-examining information, so that one can build a body of evidence towards support of an argument. The investigators regularly held peer checks with each other to identify trends, compare thematic categories and frequencies, and negotiate categorizations and code definitions. The third stage of coding regrouped the interview data per the four categories of the SAMR Model, additionally cross-referencing the content filed under these codes with the themes from the previous wave of coding. (See Tables 3, 4, and 5.)

Table 3. *Codebook I, Part I*

Codes and Subcodes	Definitions
Technology limitations	This is a general code for any discussion related to issues implementing and sustaining technology usage for teaching and learning. It can include challenges related to access and bandwidth, affordances of the technologies and their conductivity towards desired learning design, and troubleshooting.
Logistics	This subcode captures more nuanced discussion regarding challenges that the instructors experienced in coordinating classroom operations, such as learning materials, assessment, grading, and interaction in the online environment.
Software constraints	This subcode refers to instances where technology specifically limited the instructors in their instructional methods and how they wanted to deliver and coordinate activities, particularly where they ultimately could not deliver what it was they wanted to achieve. The subcode connotes when technology posed a new barrier to in-person teaching.
Things not working right away	This subcode refers to the inability of instructors to troubleshoot or receive adequate support using the technology, particularly if a learning curve curtailed its continued usage or impacted usage in any way and was abandoned despite more long-term potential.
Measures to control cheating	Strategies that instructors took up to curtail cheating whether by proactive or reactive design. Examples include proctoring or plagiarism software, using timed and multiple versions of exams, creating project-based assignments, etc.
Cheating occurrences	This subcode refers to specific, quantitative instances when instructors experienced and had to mitigate cheating in their classes. These do not refer to anything perceived or preventative.

Table 4. *Codebook I, Part II*

Codes and Subcodes	Definitions
Rigor (seatbelts)	This code refers to the supports that instructors used to modify workload, cognitive load, and wellness for their blended and/or distance learners. It may relate to a lowering of standards, offering flexibility and removing deadlines, etc. It was called "seatbelts" due to one instructor mentioning that students' mastery of his content had life or death consequences since they would eventually design safety equipment like seatbelts.
Weighting grades	This subcode refers to specific, quantitative instances when instructors changed the weighting of their assignments and grading compared to teaching in-person courses.
Flexibility and choice	This subcode refers to instances where instructors offered students more flexibility and choice regarding assignments and their deadlines, including adapting the courses based on student preferences indicated during formative evaluation.
Lower stakes exams and "chunkification"	This subcode refers to instances when instructors adapted large summative exams to more lower stakes assessment, such as weekly quizzes. It also includes examples when instructors streamlined or "chunkified" content for better cognitive processing and manageability.
Successes	This code captures high-level discussion on positive teaching experiences vis-à-vis when new strategies and tools ultimately did produce the desired outcomes.
Assessment schedule	This subcode refers to instances where instructors expressed satisfaction in adapting assessment schedules and the residual learning and performance improvements experienced by students as the result of more formative or small stakes exams. Some content may refer to the redistribution of assignments and changed weighting or timing.
Beyond memorization	This subcode refers to instances where instructors were able to use technology to create more generative learning exchanges and transcend the use of rote knowledge.
Failures	This code captures high-level discussion of negative teaching experiences vis-à-vis when new strategies and tools ultimately did not produce the outcomes desired.
Lack of participation and engagement	This subcode refers to instructors' experiences struggling to achieve participation, collaboration, engagement, and social presence in the adapted classroom.
Workload	This subcode refers to instances where instructors expressed a significant increase in the level of effort and difficulties in managing their own workload during online teaching (e.g., grading, advising students, designing and sharing content, troubleshooting, etc.)

Table 5. *Codebook II*

Codes	Definitions
Substitution	Technology acts as a direct tool substitute with no functional change. For example, replacing a lecture with a Zoom lecture or recorded video.
Augmentation	Technology acts as a direct tool substitute with functional improvement. For example, quizzes that are built into the middle of an online lecture.
Modification	Technology allows for significant task redesign. For example, utilizing a wiki or Google Doc that allows users to collaborate both asynchronously or synchronously.
Redefinition	Technology allows for the creation of new tasks previously inconceivable. For example, creating a generative, multimedia project that would not be possible in a traditional classroom.

Results

RQ 1: Identifying Technology-Enhanced Assessment Strategies Used by Instructors

The 13 instructors used a wide-variety of technology-enhanced assessment strategies in their large undergraduate classes during the Fall 2020 semester of the COVID-19 pandemic. These strategies were in addition to other approaches, such as redistributing the weighting of assignments and grades, offering more lower stakes exams and streamlining or "chunking" lessons, and giving students more flexibility and choice in terms of assignments and deadlines.

Common technology-enhanced strategies included making use of software platforms, such as Panopto, SmartBook, Packback, Gradescope, and HonorLock. However, Gradescope and HonorLock were not typically met with logistical success, being noted as "cumbersome" and "time-consuming." Turning Technologies ResponseCard® RF Clicker helped to improve student engagement during class, especially in light of university-based web camera policies that allowed students to opt out of camera usage at any time.

Videos with embedded quizzes were a largely used and successful strategy. For example,one participant commented:

> I learned how to use the Panopto recordings and how to transfer things from one place to the other and how to embed quizzes and stuff in them so that I could do assignments based on the quizzes … . I have quizzes embedded in the videos that they have to do. They're very easy, and they can basically do them until they get 100%. So, it's mostly to make sure they watch them.
>
> —Instructor, College of Behavioral and Social Sciences

All instructors using video-based quizzes remarked on their utility as a study aid for students, believing that their usage created better mental enforcement and retention of lecture material. This led to secondary benefits, such as study tools for students and lessened cognitive load. Those who incorporated video-based quizzes all noted that student performance on other larger stakes exams was higher than in years' prior without these quizzes. A visual example of theme counts and frequencies from Codebook I is included in Table 6.

Table 6. *Theme Counts and Frequency[1] for Codebook I*

Themes and Subthemes	Interview Count	Interview Frequency	Frequency Rating
Technology limitations	10	49	4.9
Logistics	8	15	1.9
Software constraints	7	12	1.7
Things not working right away	8	9	1.1
Measures to control cheating	6	13	2.2
Cheating occurrences	5	5	1.0
Rigor (seatbelts)	10	60	6.0
Weighting grades	4	7	1.8
Flexibility and choice	5	8	1.6
Lower stakes exams and "chunkification"	10	26	2.6
Student performance	7	10	1.4
Decreased performance	3	5	1.7
Increased performance	5	8	1.6
TA experiences	4	5	1.3
Successes	5	6	1.2
Assessment schedule	3	4	1.3
Beyond memorization	1	1	1.0
Failures	12	20	1.7
Lack of participation and engagement	5	9	1.8
Workload	6	10	1.7

Additionally, most of the embedded quizzes described by participants were driven by AI. This meant that they were constructed through software programs that used adaptive algorithms to generate individualized questions. Different questions would appear before students based on the content that was immediately covered, as well as their personal user answers. If a student answered incorrectly, consequent questions would be automatically generated to get easier or try to re-ask the same content in a different way. Correct answers

[1] Interview Count refers to the number of individual interviews that reference a given theme. Each theme was counted only once per interview with a maximum count reflecting the 13 interviews in total. For example, a score of 5 means that 5 interviews out of 13 discussed that theme. Interview Frequency refers to the amount of times a theme was discussed in total by interviewees even if that was multiple times within the same interview. The frequency rating was calculated by dividing the Interview Frequency by Interview Count to better indicate theme prominence.

may have led to harder questions or a follow-up question on different content. The primary purpose was not to assess learning per se but prompt engagement and active listening during the video activities. However, instructors observed that these tools did, in fact, promote content retention and cognitive processing. They also appreciated the freedom of the video as a lecture substitute and how it allowed them to bring more flipped-based approaches to their teaching. By using asynchronous video instead of in-person lecture, instructors had increased opportunities during synchronous class time to implement more engaging, innovative, and collaborative activities. This relieved a standard lecture hall monotony for both instructors and students. In the words of one interviewee:

> The post-COVID course will not look like the pre-COVID course. Students will be watching lecture videos and in class, which is really the valuable time to interact with the instructor, particularly because they, for the most part, don't want to go to office hours. This is my time, forced me to see what they're doing and forced them to perform in front of me. So, I'm looking forward to it. It's going to make lecturing a lot more fun. I'm not going to have to be up there and [do] boring lecture. I can do what I like, which is to go talk to people and, hopefully, convey a sense of enthusiasm.
>
> – Instructor, A. James Clark School of Engineering

RQ 2: Relationships Between Assessment Strategies and the SAMR Model

Across the 13 interviews, not one instructor used a singular approach. There were several mixed components in classes that relied on either Substitution, Augmentation, Modification, or Redefinition methods. As may be assumed during the forced transition to online teaching, most strategies fell on the "enhancement" versus "transformation" end of the technology integration spectrum. This is expected and has some benefits, especially towards continuity during an emergency. All 13 instructors surveyed used some form of substitution method in their teaching; for example, using Zoom to provide synchronous lectures that would normally have been delivered in a classroom. However, most Substitution attempts ended in frustration or failure, particularly due to technology constraints. For example, instructors who tried to replicate an exercise as they would do in person expressed disappointment with their student performances and marked frustrations over the loss of class interactivity, hands-on learning, and informal touchpoints for them to gather formative feedback.

Moreover, 11 out of 13 instructors used the funding to elevate other facets of their course, also utilizing Augmentation strategies. An additional visual representation of thematic counts and frequencies for Codebook II is included in Table 7. Most examples involved using technology tools to build in more study spaces and exercises, videos, formative quizzes, and discussion activities. Those utilizing Augmentation strategies in some form expressed the highest positive outcomes and experiences from the course. These improvements include increases in student focus and motivation, performance and quality of work, and overall student satisfaction. Both Substitution and Augmentation approaches hold the potential to improve teaching in blended environments and can also readily carry over into in-person teaching.

Table 7. *Theme Counts and Frequency for Codebook II*

Themes	Interview Count	Interview Frequency	Frequency Rating
Substitution	13	26	2.0
Augmentation	11	21	1.9
Modification	8	12	1.5
Redefinition	1	1	1.0

Additionally, eight out of 13 instructors worked to Modify lessons. However, these strategies received mixed outcomes, possibly due to low beta testing, and they were more likely to create additional workload for instructors. Only one instructor created an assignment that allowed them to fully redefine an activity. They state:

> I have actually changed some of my setup. For example, next Monday, students can book time with me … time outside our regular class. [Throughout the day] student teams can log in into the main class but, then, can jump into their private room. At a specific time, I'll show up. I have 10 minutes with four students in the team project, and it's absolutely something I cannot do in person. It just doesn't work. Here, I have a completely safe environment where students can talk about their project, and I'm able to help them move the project along.
>
> —Instructor, Robert H. Smith School of Business

However, having only one case of Redefinition indicates that this was not a common approach to remote teaching during the COVID-19 Pandemic.

RQ 3: Reflections on the Benefits and Challenges of Online Assessment Strategies

Five out of the 13 interviewed instructors largely felt that the online learning experience had a positive impact on improving learning outcomes and their students' overall performances. In contrast, three instructors relayed a marked decrease in student performance. Those with positive outcomes also mutually shared the sentiment that these were unanticipated results:

> Right now, I'm so surprised. Usually, [I] have at least 25–35% of the students based on the past 5 years in the failing category at this point. Right now, we have only one … . I'm pretty pleased. That's despite the COVID … If the COVID was not here, who knows? It could have been even better.
>
> —Instructor, A. James Clark School of Engineering

To further clarify, the projection that the results of online learning would have been higher if not for the COVID-19 pandemic was in acknowledgement of socio-emotional wellbeing during these times—the idea being that students still managed to perform at better rates despite assumed strains in physical and mental health. In line with prior literature, the instructors also struggled with the amount of work required for online course design, whether through planning, implementation, or grading:

> I felt like with this course I painted myself into the corner with how much grading there has been. It has been nonstop grading. It's been a marathon of grading 7 days a week.
>
> —Instructor, Philip Merrill College of Journalism

The burden of the workload was connected to concerns over effective proctoring and academic dishonesty. Many instructors discussed strategies for better grading efficiency and managing their workloads. Two participants had great success with introducing peer review as part of the assessment process. Specifically, they removed or significantly reduced traditional exams and developed assignments as the larger portion of the final grade. In turn, these instructors claimed that the peer review process had a significant positive impact on student performance and engagement in class, being seen by those who implemented it as one of the most successful assessment strategies:

> We reduced the percent by 7.5% that we counted towards exams, and we put it into this Packback [web platform]. The students' peer-to-peer discussions, helping them make connection to the course material, helping them to understand intuitively what they were doing rather than just memorizing

> steps. It was due every Friday at 11 o'clock. That was one part of this [online teaching experience] I think was very much a success.
>
> —Instructor, Robert H. Smith School of Business

Student grading of assignments improved interactivity and engagement, both of which were challenging in online large lecture-style courses. It also helped to partially mitigate hands-on learning loss and scaffold activities that spoke to more learner-centered pedagogy—all while creating a lighter-touch review for final grading by the instructor.

Multiple instructors (n = 5) had the perception that academic dishonesty would be a major challenge, taking proactive steps to try and circumvent it through design. However, only one instructor out of the 13 experienced overt issues with student cheating. At the same time, no participant described success using proctoring software, particularly Honorlock, and discussed many frustrations, limitations, and/or technical difficulties, as mentioned.

Overall, a common approach was instructors' use of the Innovation Grant funding to develop multiple versions of exam questions and layout over the summer. In some ways, this impacted instructor workload during the design phase but alleviated the need to do further development mid-course. Three participants spoke about the tasks of creating multiple test versions, including asking more critical and problem-oriented questions so as to purposefully prevent the ability of students to readily look up answers. This prevention was based on the nature of the questions' content as well as instructors' assumptions that such questions would be too time consuming to search online. Another common tactic was to time these assessments to prevent such cheating—or just make the final exams open-book but with similar design parameters. These measures are considered a Substitution strategy.

Discussion

Integration Into the Current Literature

The findings demonstrate several ways in which technology-enhanced assessment strategies were implemented in large undergraduate classes during the Fall 2020 semester of the COVID-19 pandemic. On a practical level, instructors followed many of the approaches recommended in the literature for standard online teaching (i.e., take-home tests, scrambling questions/multiple versions of exams, designing questions in such a way as to be hard to look up, etc.) and experienced marked success in student performance. Although some assessment policies did vary from the standard literature due to additional emergency response and wellness considerations, the four participants who used more holistic approaches indicated that they saw such marked improvement to student performance that they would continue to develop and use them in their classes for the future, regardless of pandemic or emergency response teaching. In our sample, more instructors put an emphasis on an honor code, choosing to relax hard submission deadlines and offer other flexibilities to support mental health and mitigate stress during the pandemic. These considerations were in line with the evidence that stress, vulnerability, and high stakes assessment all increase academic misconduct (Eaton, 2020; Nichols & Berliner, 2007). Although participants invested substantial time and resources per the grant funding into reconfiguring assessment design to get ahead of potential academic dishonesty, the five instructors who found ways to incorporate more formative milestones and student feedback during the semester indicated the highest student performance. None found success when trying to use formal proctoring software (Substitution), and most found even greater challenges with student motivation and burnout. In these instances, more student autonomy/flexibility and reflection in assessment appears to have successfully mitigated cheating—with only one known case occurring within our sample.

However, instructors reflected on how these measures negatively impacted their workloads, citing workload as the primary challenge regarding their assessment strategies—and that the inefficiency was not sustainable—"a marathon of grading 7 days a week." It took more time for instructors to grade writing-based assignments, as well as more frequent lower-stakes quizzes, certainly complicated by their large class sizes. For those who took this approach, several other challenges were cited, such as developing intricately coordinated exam versions, question sequences, grading keys, algorithms, and problem-centric prompts. It is little wonder that the instructors who found ways to implement a peer review process found two-fold success: peer review helped to alleviate the grading workload of the teachers while simultaneously fostering student engagement and learner-centered pedagogy. Instructors whose assessments were centered on peer-based assessment strategies and/or formative feedback also indicated better student performance than in their previous pre-COVID teaching. Per the SAMR Model, these two approaches operated as higher-order strategies, transforming the remote assessment process rather than just enhancing it. Instead of substituting traditional exams with some form of electronic delivery or even lower stakes quizzes, instructors were able to transform lessons through the electronically facilitated means of participatory evaluation and digital authorship, both approaches backed by the empirical evidence on the power of student reflection (Black & Wiliam, 1998; Broadfoot et al., 2001; Gikandi et al., 2011; Shepard, 2000; Timmis et al., 2015). For these large classes, instructors needed to find ways to break the material into smaller work groups and units. Per the literature, this both alleviates the higher workload associated with online teaching, as well as puts less burden on the higher cognitive load that online reading requires (CAST, 2018), fueling screen fatigue and diminishing retention. Here, instructors indicated that peer review helped maximize learning and promote efficiency. The peer review strategy was also in line with the SAMR Model—rather than converting a print-based, closed-ended exam into a digital format as a form of Substitution, peer review allowed for more collaborative and inquiry-based reflection, Modifying or Redefining the assessment and overall learning experiences.

Limitations and Implications for Theory and Practice

Although we have taken these results from a small sample, they do show some key trends and help to further explore the evidence base on what works for online summative assessment, particularly with regards to how to administer assessment in large-size, lecture-style undergraduate classrooms—the type often experienced at public research universities. The COVID-19 pandemic may have been the overall catalyst for course adaptation, but the residual impact of the pandemic on classrooms can only provide a barometer for how well adaptations and performance will continue to fare under other emergencies and disruptions.

As a limitation, the researchers did not have access to student final grades or student perspectives on the assessments and overall course experiences. Future research should work to address this gap, as well as review which assessment practices were maintained by the instructors for future sessions, especially as the pandemic wanes and a "new normal" emerges. To this end, there will be a need for further investigation into scale-up and sustainability of these models over long-term course delivery. Our findings show much potential for the post-coronial university, offering three main takeaways:

1. **Academic dishonesty is diminished by online assessment**, especially when the mediated classroom allows for more student autonomy, creative work, and deadline flexibility.
2. **Online assessment offers multiple benefits.** Digital affordances allow instructors to take on new opportunities for critical evaluation and maker education, including interactive peer review and increased touchpoints for formative assessment. These do not need to be excluded from large-sized classes, at least when designed in ways to diminish instructor workload.
3. **Innovation requires investment.** Without the funding to cover time and materials for course redesign, instructors would have been unable to develop promising, alternative assessment approaches. This is indicated by the workload that was experienced by instructors at the course-design level rather than at the student-assessment level.

Conclusion

The goal of this exploratory research was to better understand the types of assessment strategies employed by the University of Maryland, College Park instructors who received an Innovation Grant from the Provost's Office in Summer 2020. Despite the three most recurrent themes being technology limitations, failures, and rigor (seatbelts) (i.e., the supports that instructors used to modify workload, cognitive load, and wellness for their blended and/or distance learners), instructors experienced several successes during course retooling, including significant increases in student performance. These were attributed to new forms of assessment used for improving student engagement and formative feedback—such as the use of video-embedded CAT or peer review. The results showed that instructors used a mix of practices under SAMR but most often relied on Substitution or Augmentation strategies. While most Substitution attempts ended in frustration or failure, those applying Augmentation strategies expressed the highest positive student outcomes and teaching experiences overall. These improvements included increases in student focus and motivation, performance and quality of work, and overall student satisfaction. Most instructors indicated that they would continue to keep these new online assessment strategies and much of their course changes for the future, whether that future includes blended or in-person delivery.

Surprisingly, despite the anticipation that the pandemic would exacerbate student stress and that online teaching would fuel more opportunities for cheating, instructors found few perpetuations of academic dishonesty. Simply, instructors provided more understanding, flexible deadlines, and empathy, which led to less cheating. In some ways, the key to improving academic integrity in assessment may be proactive approaches to test design and less policing, not more. Further investigation into this phenomenon would be beneficial for the future delivery of higher education, particularly in view of initiatives for improving both student wellness and retention. New research would also benefit from gathering student and teaching assistant voices on these learning experiences beyond the instructors' perspectives.

References

Attwood, R., & Radnofsky, L. (2007, Sept. 14). Satisfied—but students want more feedback. *Times Higher Education*. https://www.timeshighereducation.com/news/satisfied-but-students-want-more-feedback/310440

Baron, J., & Crooks, S. M. (2005). Academic integrity in web based distance education. *TechTrends, 49*(2): 40–45.

Barros, A., & Joung, M. (2020, March 10). Colleges, universities move classes online amid coronavirus outbreak. *VOA*. https://www.voanews.com/a/science-health_coronavirus-outbreak_colleges-universities-move-classes-online-amid-coronavirus/6185570.html

Bartoletti, R. (2016). Learning through design: MOOC development as a method for exploring teaching methods. *Current Issues in Emerging eLearning, 3*(1), Article 2. https://scholarworks.umb.edu/ciee/vol3/iss1/2

Bawa, P. (2016). Retention in online courses: Exploring issues and solutions—a literature review. *SAGE Open*. https://doi.org/10.1177/2158244015621777

Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy and Practice, 5*(1): 7–74. https://www.gla.ac.uk/t4/learningandteaching/files/PGCTHE/BlackandWiliam1998.pdf

Bonk, C. J., Zhu, M., Kim, M., Xu, S., Sabir, N., Sari, A. R. (2018). Pushing toward a more personalized MOOC: Exploring instructor selected activities, resources, and technologies for MOOC design and implementation. *International Review of Research in Open and Distributed Learning, 19*(4), 93–115. https://www.irrodl.org/index.php/irrodl/article/view/3439/4765

Boud, D., & Molloy, E. (Eds). (2013). *Feedback in higher and professional education: Understanding it and doing it well*. Routledge.

Broadfoot, P., Osborn, M., Sharpe, K., & Planel, C. (2001). Pupil assessment and classroom culture: A comparative study of the language of assessment in England and France. In D. Scott (Ed.), *Curriculum and assessment*. Ablex.

Brouwer, N., & McDonnell, C. (2009). Online support and online assessment for teaching and learning chemistry (pp. 123–152). In I. Eilks & B. Byers (Eds.), *Innovative methods of teaching and learning chemistry in higher education*. RSC.

Buckle, J. (n.d.). *A comprehensive guide to 21st century skills*. Panorama Education. https://www.panoramaed.com/blog/comprehensive-guide-21st-century-skills

Care, E., & Vista, A. (2017, February 23). *Education assessment in the 21st century: Moving beyond traditional methods*. Brooking Institution. https://www.brookings.edu/blog/education-plus-development/2017/02/23/education-assessment-in-the-21st-century-moving-beyond-traditional-methods/

Carnevale, D. (1999). How to proctor from a distance. *Chronicle of Higher Education, 46*(12), A47–A48.

CAST. (2018). *Universal Design for Learning guidelines*. http://udlguidelines.cast.org

Cluskey, G. R., Jr., Ehlen, C. R., & Raiborn, M. H. (2011). Thwarting online exam cheating with proctor supervision. *Journal of Academic and Business Ethics, 4*. https://www.aabri.com/manuscripts/11775.pdf

Conrad, D., & Openo, J. (2018). *Assessment strategies for online learning: Engagement and authenticity*. AU Press.

Costa, S. A., Kavouras, I., Cohen, N., & Huang, T. T. K. (2021, October 25). Moving education online during the COVID-19 pandemic: Thinking back and looking ahead. *Frontiers in Public Health, 9*. https://doi.org/10.3389/fpubh.2021.751685

Crisp, V., & Ward, C. (2008). The development of a formative scenario-based computer assisted assessment tool in psychology for teachers: the PePCAA project. *Computers & Education, 50*(4), 1509–1526. https://doi.org/10.1016/j.compedu.2007.02.004

Davis, S. F., Drinan, P. F., & Bertram Gallant, T. (2009). *Cheating in school: What we know and what we can do*. Wiley-Blackwell.

Deneen, C., & Boud, D. (2014). Patterns of resistance in managing assessment change. *Assessment & Evaluation in Higher Education, 39*(5), 577–591. https://doi.org/10.1080/02602938.2013.859654

Eaton, S. E. (2020). Academic integrity during COVID-19: Reflections from the University of Calgary. *International Studies in Educational Administration, 48*(1): 80–85. https://prism.ucalgary.ca/server/api/core/bitstreams/40b2bf53-b117-4e4d-a358-02aa7fa7b460/content

Elsalem, L., Al-Azzam, N., Jum'ah, A. A., & Obeidat, N. (2021). Remote e-exams during Covid-19 pandemic: A cross-sectional study of students' preferences and academic dishonesty in faculties of medical sciences. *Annals of Medicine and Surgery,* 62, 326–333. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7825891/

Elwood, J., & Klenowski, V. (2002). Creating communities of shared practice: The challenges of assessment use in learning and teaching. *Assessment & Evaluation in Higher Education, 27*(3): 243–256. https://doi.org/10.1080/02602930220138606

Gamage, K. A. A., de Silva, E. K., & Gunawardhana, N. (2020). Online delivery and assessment during COVID-19: Safeguarding academic integrity. *Education Sciences, 10*. https://www.mdpi.com/2227-7102/10/11/301

Gaytan, J., & McEwen, B. C. (2007). Effective online instructional and assessment strategies. *American Journal of Distance Education, 21*(3), 117–132. https://doi.org/10.1080/08923640701341653

Gee, J. P., & Shaffer, D. W. (2010). Looking where the light is bad: Video games and the future of assessment. *Phi Delta Kappa International EDge, 6*(1), 3–19. https://www.researchgate.net/publication/303856565_Looking_where_the_light_is_bad_Video_games_and_the_future_of_assessment

Gikandi, J. W., Morrow, D., & Davis, N. E. (2011). Online formative assessment in higher education: A review of the literature. *Computers & Education, 57,* 2333–2351. https://doi.org/10.1016/j.compedu.2011.06.004

González-Pérez, L. I., & Ramírez-Montoya, M. S. (2022). Components of education 4.0 in 21st century skills frameworks: Systematic review. *Sustainability*, *14*(3), 1493. https://doi.org/10.3390/su14031493

Guzmán, E., & Conejo, R. (2005). Self-assessment in a feasible, adaptive web-based testing system. *IEEE Transactions on Education, 48*(4), 688–695. https://www.researchgate.net/publication/3051034_Self-Assessment_in_a_Feasible_Adaptive_Web-Based_Testing_System

Hart, L., & Morgan, L. (2010). Academic integrity in an online registered nurse to baccalaureate in nursing program. *The Journal of Continuing Education in Nursing, 41*(11), 498–505.

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research, 77*(1), 81–112.

Heberling, M. (2002). Maintaining academic integrity in online education. *Online Journal of Distance Learning Administration, 5*(1). http://www.westga.edu/%7Edistance/ojdla/spring51/heberling51.html

Hone, K. S., & El Said, G. R. (2016). Exploring the factors affecting MOOC retention: A survey study. *Computers & Education, 98,* 157–168. https://doi.org/10.1016/j.compedu.2016.03.016

INEE. (2018). INEE Guidance note on psychosocial support: Facilitating psychosocial wellbeing and social and emotional learning. https://app.mhpss.net/?get=211/inee_guidance_note_on_psychosocial_support_eng.pdf

Jian-quan, T., Dan-min, M., Xia, Z., & Jing-jing, G. (2007). An introduction to the Computerized Adaptive Testing. *US-China Education Review, 4*9(1), 72–81. https://files.eric.ed.gov/fulltext/ED497385.pdf

Kamei, A., & Harriott, W. (2021). Social emotional learning in virtual settings: Intervention strategies. *International Electronic Journal of Elementary Education*, *13*(3), 365–371. https://www.iejee.com/index.php/IEJEE/article/view/1487

Katzman, N. F., & Stanton, M. P. (2020). The integration of social emotional learning and cultural education into online distance learning curricula: Now imperative during the COVID-19 pandemic. *Creative Education, 11*(9), 1561–1571. https://doi.org/10.4236/ce.2020.119114

Koh, L. C. (2008). Re-focusing formative feedback to enhance learning in pre-registration nurse education. *Nurse Education in Practice, 8*(4), 223–230. https://doi.org/10.1016/j.nepr.2007.08.002

Korkmaz, G., & Toraman, Ç. (2020). Are we ready for the post-COVID-19 educational practice? An investigation into what educators think as to online learning. *International Journal of Technology in Education and Science, 4,* 293–309. https://eric.ed.gov/?id=EJ1271308

Lin, Q. (2008). Preservice teachers' learning experiences of constructing e-portfolios online. *Internet and Higher Education, 11*(3), 194–200.

Looney, J. W. (2011), *Integrating formative and summative assessment: Progress toward a seamless system?* OECD Education Working Papers, No. 58. http://dx.doi.org/10.1787/5kghx3kbl734-en

Ludwig-Hardman, S., & Dunlap, J. C. (2003). Learner support services for online students: Scaffolding for success. *International Review of Research in Open & Distance Learning, 4*(1), 1–15. https://www.researchgate.net/publication/26468303_Learner_Support_Services_for_Online_Students_Scaffolding_for_success/link/57a4d51408ae3f45292d007e/download

Mackey, J. (2009). Virtual learning and real communities: Online professional development for teachers. In E. Stacey, & P. Gerbic (Eds.), *Effective blended learning practices: Evidence-based perspectives in ICT-facilitated education* (pp. 163–181). IGI Group. https://www.researchgate.net/publication/234762996_Effective_Blended_Learning_Practices_Evidence-Based_Perspectives_in_ICT-Facilitated_Education

MacQueen, K. M., McLellan-Lemal, E., Bartholow, K., & Milstein, B. (2008). Team-based codebook development: Structure, process, and agreement. In G. Guest & K. M. MacQueen (Eds.), *Handbook for team-based qualitative research* (pp. 119–135). AltaMira Press.

Meyer, J. P., & Zhu, S. (2013). Fair and equitable measurement of student learning in MOOCs: An introduction to item response theory, scale linking, and score equating. *Research & Practice in Assessment, 8*, 26–39. https://www.rpajournal.com/dev/wp-content/uploads/2013/05/SF3.pdf

Namey, E., Guest, G., Thairu, L., & Johnson, L. (2008). Data reduction techniques for large qualitative data sets. In G. Guest & K. M. MacQueen (Eds.), *Handbook for team-based qualitative research* (pp. 137–161). AltaMira Press.

Nichols, S. L., & Berliner, D. C. (2007). *Collateral damage: How high-stakes testing corrupts America's schools* (2nd ed.). Harvard Education Press.

Nicol, D., & Macfarlene-Dick, D. (2006). Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. *Studies in Higher Education, 31*(2), 199–218. https://doi.org/10.1080/03075070600572090

Olsen, B. (2021). *Scaling education innovations for impact in low- and middle-income countries during COVID: Reflections on key themes* [Policy Brief]. Center for Universal Education at Brookings. https://www.brookings.edu/wp-content/uploads/2021/12/ROSIE-reflection-brief_FINAL-1.pdf

Oosterhof, A., Conrad, R. M., & Ely, D. P. (2008). *Assessing learners online.* Pearson.

Oppl, S., Reisinger, F., Eckmaier, A., & Helm, C. (2017). A flexible online platform for computerized adaptive testing. *International Journal of Educational Technology in Higher Education, 14*(2). https://doi.org/10.1186/s41239-017-0039-0

Puentedura, R. R. (2006, August 18). *Transformation, technology, and education*. Hippasus. Strengthening Your District Through Technology workshops, coordinated by the Maine School Superintendents Association. http://hippasus.com/resources/tte/puentedura_tte.pdf

Rajet, S., & Stitzel, S. (2020). Strategies for effective assessments while ensuring academic integrity in general chemistry courses during COVID-19. *Journal of Chemical Education, 97,* 3426–3440. https://doi.org/10.1021/acs.jchemed.0c00797

Rapanta, C., Botturi, L., Goodyear, P., Guàrdia, L., & Koole, M. (2020). Online university teaching during and after the Covid-19 crisis: Refocusing teacher presence and learning activity. *Postdigital Science and Education, 2,* 923–945. https://link.springer.com/article/10.1007/s42438-020-00155-y

Salas-Morera, L., Arauzo-Azofra, A., & García-Hernández, L. (2012). Analysis of online quizzes as a teaching and assessment tool. *Journal of Technology and Science Education, 2*(1): 39–45. https://www.jotse.org/index.php/jotse/article/view/30/52

Salcedo, P., Pinninghoff, M. A., & Contreras, R. (2005). Computerized adaptive tests and item response theory on a distance education platform. In J. Mira & J. R. Álvarez (Eds.), *Artificial intelligence and knowledge engineering applications: A bioinspired approach* (pp. 613–621). Springer.

Saldaña, J. (2016). *The coding manual for qualitative researchers*. (3rd ed.). Sage.

Shepard, L. (2000). The role of assessment in a learning culture. *Educational Researcher*, *29*(7), 1–14. https://doi.org/10.3102/0013189X029007004

Shute, V. J., & Kim, Y. J. (2014). Formative and stealth assessment. In J. M. Spector, M. D. Merrill, J. Elen, & M. J. Bishop (Eds.), *Handbook of research on educational communications and technology* (4th ed.). (pp. 311–323). Springer.

Sorensen, E. K., & Takle, E. S. (2005). Investigating knowledge building dialogues in networked communities of practice: A collaborative learning endeavor across cultures. *Interactive Educational Multimedia, 10*, 50–60. https://dr.lib.iastate.edu/server/api/core/bitstreams/06d8c740-18e3-4311-b364-f8a6d1abf2db/content

Stuber-McEwen, D., Wiseley, P., & Hoggatt, S. (2009). Point, click, and cheat: Frequency and type of academic dishonesty in the virtual classroom. *Online Journal of Distance Learning Administration, 12*(3). https://www.westga.edu/~distance/ojdla/fall123/stuber123.html

Timmis, S., Broadfoot, P., Sutherland, R., & Oldfield, A. (2016). Rethinking assessment in a digital age: Opportunities, challenges and risks. *British Educational Research Journal, 42*(3), 454–476. https://research-information.bris.ac.uk/ws/portalfiles/portal/90342446/TEA_paper_revised_following_review_Berj_final3.pdf

Tunc, E., Hossain, N., Haq, A., & Yildirim, Y. (2022). Communicating social and emotional needs: New challenges for education during the COVID-19 pandemic. In V. Trif (Ed.), *Basic communication and assessment prerequisites for the new normal of education*. (pp. 162–170). IGI Global.

Vonderwell, S., Liang, X., & Alderman, K. (2007). Asynchronous discussions and assessment in online learning. *Journal of Research on Technology in Education, 3*(3): 309–328. https://doi.org/10.1080/15391523.2007.10782485

Wang, T., Wang, K., & Huang, S. (2008). Designing a web-based assessment environment for improving pre-service teacher assessment literacy. *Computers & Education, 51*(1), 448–462. https://doi.org/10.1016/j.compedu.2007.06.010

Webb, M., & Gibson, D. (2015). Technology enhanced assessment in complex collaborative settings. *Education and Information Technologies, 20*(4), 675–695. https://doi.org/10.1007/s10639-015-9413-5

Williams, K. M., & Corwith, A. (2021). Beyond bricks and mortar: The efficacy of online learning and community-building at College Park Academy during the COVID-19 pandemic. *Education and Information Technologie*s, *26*(5), 5055–5076. https://doi.org/10.1007/s10639-021-10516-0

Witze, A. (2020, June 1). Universities will never be the same after the coronavirus crisis. *Nature*. https://www.nature.com/articles/d41586-020-01518-y

Wolcott, H. F. (1994). *Transforming qualitative data: Description, analysis, and interpretation*. Sage.

Wolsey, T. (2008). Efficacy of instructor feedback on written work in an online program. *International Journal on eLearning, 7*(2), 311–329.

Wyatt-Smith, C., & Kimber, K. (2009). Working multimodally: Challenges for assessment. *English Teaching: Practice and Critique, 8*(3), 70–90. https://files.eric.ed.gov/fulltext/EJ869395.pdf

Zhu, M., Bonk, C. J., & Sari, A. R. (2018). Instructor experiences designing MOOCs in higher education: Pedagogical, resource, and logistical considerations and challenges. *Online Learning, 22*(4), 203–241. https://doi.org/10.24059/olj.v22i4.1495

Zieher, A. K., Cipriano, C., Meyer, J. L., & Strambler, M. J. (2021). Educators' implementation and use of social and emotional learning early in the COVID-19 pandemic. *School Psychology, 36*(5), 388–397. https://doi.org/10.1037/spq0000461

JERAP

The *Journal of Educational Research and Practice* is a peer-reviewed journal that provides a forum for studies and dialogue about developments and change in the field of education and learning. The journal includes research and related content that examine current relevant educational issues and processes. The aim is to provide readers with knowledge and with strategies to use that knowledge in educational or learning environments. *JERAP* focuses on education at all levels and in any setting, and includes peer-reviewed research reports, commentaries, book reviews, interviews of prominent individuals, and reports about educational practice. The journal is sponsored by The Richard W. Riley College of Education and Human Sciences at Walden University, and publication in *JERAP* is always free to authors and readers.