A pedagogical evaluation of an institution's digital assessment platform (DAP): Integrating pedagogical, technical and contextual factors

Christine Slade, Kathleen Mahon, Karen Benson, Jacqui Lynagh, Dom McGrath, Karen Sheppard
Institute for Teaching and Learning Innovation, The University of Queensland, Brisbane, Australia

Qasim Ahsan
Independent researcher

An assessment for learning approach is foundational for student learning. The necessity to shift teaching and learning online as a response to COVID-19 has propelled digital assessment into the mainstream within higher education institutions. User experience is a common indicator of effectiveness of technologically enhanced initiatives; however, measuring the pedagogical impact of digital assessment initiatives has not been straightforward. This paper presents a pedagogical evaluation, incorporating staff and student perspectives, of a digital assessment platform (DAP) at a large Australian metropolitan university. Data were collected using surveys, interviews, focus groups and case studies and incorporated findings from student evaluations of courses and a student-staff partnership project. Results highlight strong pedagogical reasons for continuing the digital assessment initiative, while also identifying constraints and opportunities for improvement. The insights generated through this evaluation emphasise the importance of considering technical, pedagogical and contextual factors and the interplay between these factors. In a quality assurance climate, where the demand for this type of institution-wide evaluation is increasing, this study contributes an approach that addresses some of the complexities and challenges of evaluating digital assessment initiatives in higher education. Future research is needed to understand more fully the relationships between pedagogical and contextual factors when undertaking an evaluation.

Implications for practice or policy:

- The DAP enabled the (re)design of authentic, inclusive, engaging and secure assessment tasks.
- Educator support is critical in learning the system and creating assessments with high pedagogical merit.
- Student familiarisation opportunities with using the DAP are beneficial.
- The DAP's longevity within the institution is crucial considering the significant effort required of educators and students.
- Contextual influences blur understanding of pedagogical factors.

Keywords: digital assessment, digital assessment platform, e-assessment, pedagogical evaluation, pedagogical and contextual indicators, institutional context

Introduction

Assessment for learning is crucial to learning processes in higher education but poses significant challenges due to the high stakes attached to assessment outcomes. Digital technologies have introduced new ways to respond to assessment challenges, offering enhanced representation of knowledge and skills, increased flexibility in time and location, enhanced collaboration, facilitated assessment of complex problem-solving skills and enhanced feedback to students (Timmis et al., 2016).

However, a gap exists between the possibilities of technology and students' user experience (Henderson et al., 2017), and, as Sweeney et al. (2017) noted, where technology is adopted, it is often used to replicate existing assessment practices rather than to transform them. Perceived barriers, such as time constraints

and increased workload, lack of confidence and inadequate technical and pedagogical skills, support and training (Mercader & Gairin, 2020; Teo, 2019), hinder adoption. Educators' underlying personal beliefs and attitudes concerning teaching, learning and technology (Celik & Yesilyurt, 2013), as well as tensions around administrative challenges (e.g., when implementing e-exams – see Fluck, 2019), may also be contributing factors.

Capitalising on affordances of digital technologies and addressing adoption challenges, a large metropolitan Australian university introduced a commercial, cloud-based digital assessment platform (DAP) to transform assessment across its faculties and schools. This paper presents a pedagogical evaluation of that initiative. Pedagogical evaluation here refers to the initiative's teaching and learning gains and challenges, such as the consequences for teaching and learning, including achievement of pedagogical goals; impact on teaching and assessment practices and/or student learning, engagement, and needs; teacher professional learning opportunities; and changes to the teaching and learning environment. The aim of the paper is to highlight the affordances and challenges of using the DAP for pedagogical benefit and how complexities of pedagogically evaluating digital assessment initiatives in higher education can be navigated.

The paper begins with a brief discussion of digital assessment and challenges regarding evaluation in the higher education context, followed by background information about the DAP and the evaluation study. Next, an outline of the methods is provided, including key pedagogical and contextual considerations that framed data collection and analysis. Key findings are then presented and discussed, first with a focus on staff perspectives and second with a focus on student perspectives. Implications for evaluating digital assessment developments within the contemporary digital and higher education landscape are considered in this discussion. The conclusion provides suggestions to guide future studies.

Digital assessment and evaluation in higher education

Given that assessment is a critical feature of higher education, it is not surprising that there are concerns with quality and effectiveness and regular calls for renewal (see Boud et al., 2010; Bearman, Boud, & Ajjawi, 2020; Gneil, 2023). Recommendations for enhancing assessment include improved alignment between the learning outcomes, the teaching and learning activities and the assessment (Wang et al., 2013); appropriate task-specific criteria and standards (Sadler, 2005); moderation and calibration (Bloxham & Boyd, 2007; Sadler, 2017); security (Dawson, 2020); timely and effective feedback (Ryan et al., 2021; Winstone & Carless, 2020); inclusion (Tai et al., 2023); opportunities for students to self-assess (Yan et al., 2023); and student partnerships (Bovill et al., 2021).

Recent technological advancements support calls for renewal and enhancement, leading to the rise of digital assessment. The term digital assessment is used in this paper to denote all assessment that leverages digital technologies, acknowledging it may be used interchangeably with online, computer-based assessment or e-assessment – all of which we regard as subsets of digital assessment. Despite challenges, such as initially increasing assessor workload (St-Onge et al., 2022), digital assessment is reported to enable more "engaging, effective, and learner-centred" (personalised) assessment methods (Ahmed & Sidiq, 2023, p. 11), possibilities for rapid, sometimes immediate, feedback (St-Onge et al., 2022) and enhanced student learning (Mimirinis, 2019) if "properly designed" (Appiah & van Tonder, 2018, p. 1458). Digital assessment platforms are also used in efforts to streamline and standardise assessment administration through automated assessment design, distribution and grading processes.

Evaluating the use of digital assessment and platforms, however, is complex. Educational impacts are difficult to measure due to a multitude of influencing factors (Chalmers & Hunt, 2016) and the difficulty in establishing causal relationships (Guskey, 2000). The complexity of assessment itself adds to the challenge, evidenced by the lack of consensus of assessment quality evaluation criteria in higher education (Gerritsen-van Leeuwenkamp et al., 2017). For instance, criteria such as validity, reliability, transparency, authenticity, consistency or alignment and educational consequences are widely recognised, but their application and interpretation vary in higher education literature (see Appiah & van Tonder, 2018; Gerritsen-van Leeuwenkamp et al., 2017; Schellekens, 2023). Further, the interplay

between technology, pedagogy and content in digital education is complex (Mishra & Koehler, 2006), arguably made more so by the rate at which technology advances. Digital tools introduce additional factors that influence impact, such as digital literacy and access to technology (Ahmed & Sidiq, 2023) and tool functionality (Davis, 1986).

Frameworks such as the technology acceptance model (TAM), and its iterations TAM2 (Venkatesh & Davis, 2000) and TAM3 (Venkatesh & Bala, 2008), provide insights on individual user experience and user behaviour regarding, a system's design features (Davis, 1986). The technological pedagogical content knowledge framework explores the interplay between technology and teaching through three overlapping areas: content knowledge, pedagogy and technology (Mishra & Koehler, 2006, p. 1019). Kemp et al. (2019) have provided a broader evaluative taxonomy for higher education, introducing the "instructor attributes" category, which recognises design factors can influence the richness of the learning experience through feedback, interaction and selected content. Although these models provide valuable insights into understanding user acceptance of technology, they have limitations in terms of the relationship between technological enablement, influencing factors and pedagogical gains, and so fall short in facilitating the pedagogical impact evaluation of digital assessment initiatives.

Huber et al. (2023, p. 214) proposed Padayachee et al.'s (2021) 4-pillar supportive online assessment framework as one that prioritises "the elements of the learning environment that support student learning and less on the technological aspects of online assessment" (p. 214). Furthermore, Huber et al. developed and evaluated a new evidence-based framework that includes six assessment design considerations (equity of access, academic integrity, student experience, authenticity, information integrity and quality feedback) and four contextual factors (accreditation, institutional policies, scale and resourcing), which resonated more closely, but not completely, to the institutional environment of the university in this study.

Despite these models, there is still a pressing need for further research and theoretical development in this field. This paper contributes to addressing this gap by situating a practice-orientated evaluation approach within current literature and offers insights into the relationships between pedagogical, technological and contextual factors that occur in such work. We acknowledge that this paper is intended as an explorative, iteratively reflexive piece based on our struggles to deal with the complexities of evaluating digital assessment developments in the absence of appropriate frameworks.

The DAP and the evaluation study

Pre-COVID-19, the university tasked an e-assessment working group to identify digital assessment options that both facilitated improved pedagogical practices and addressed existing administrative issues. A key component was the selection and implementation of a centrally supported digital assessment platform to be available to all schools and faculties as required, which aligned with most user requirements and could be integrated seamlessly into the existing institutional information technology infrastructure. Following extensive market research and procurement activities, a new cloud-based digital assessment platform, referred to as the DAP in this paper, which supported a range of assessment types, including digital on-campus examinations, was selected. In 2022, extra components that integrated with the platform to support remotely invigilated off-campus examinations were made available.

After two pilot semesters, the DAP was offered as an opt-in-opportunity via a school-based rollout from 2021. Over five semesters, approximately 23,500 students used this new platform for assessments, such as online exams, tutorial problem sets, laboratory practicals, revision tasks, theory to practice exercises, data exercises, end of rotation assessments and reflections. Just under 10% were case study assessments, reports and video assignments. Strategically, assessment design and redesign were seen as the foundation of the work conducted within this project, so teaching teams were supported by expert learning designers to redesign assessments, with a focus on contemporary ideas about assessment and feedback and the creation of authentic assessment to enhance the student experience.

In 2023, the university's central teaching and learning unit conducted a pedagogical evaluation to determine the extent to which, and how, assessment processes and outputs designed and implemented using the DAP and associated supports, had been pedagogically transformed. This work built on an initial pilot evaluation conducted in November 2021 that aimed to inform the future development of digital assessment at the university.

To evaluate the impact of using the DAP on pedagogical practices and outcomes, we focused on the experiences of key stakeholders (staff and students), guided by the following research questions:

- What are the staff members' perspectives on their engagement with, and outcomes of the DAP initiative?
- What are the students' perspectives on assessment they completed using the DAP?
- How did contextual factors impact the pedagogical outcomes of the DAP?

We drew on existing literature to identify a framework that focused on pedagogical rather than technical or administrative concerns. The Huber et al. (2023) framework, despite being developed for business education and online assessment contexts, was adopted due to its (a) applicability to digital assessment, (b) focus on student learning and the student experience, (c) evidence-informed design considerations and contextual factors and (d) alignment with the university's assessment policy which highlights commitments to authentic, purposeful, fair and equitable assessment, academic integrity and quality feedback.

Yet the framework was not entirely suited to the DAP initiative, so rather than adopting it in its entirety, we used it to develop key factors of assessment and the assessment context to guide data collection and analysis. These elements were divided into pedagogical factors (Table 1) (replacing design considerations in the original framework) and contextual factors (Table 2)

Table 1
List of pedagogical factors used in the evaluation

Pedagogical factor	Description
Academic integrity and assessment security	Academic integrity is a commitment to six fundamental values: honesty, trust, fairness, respect, responsibility and courage (International Center for Academic Integrity, 2021). Assessment security is the restrictions placed on assessment tasks to ensure, as best as possible, that the student who gets credit for the task is the same person who does the work (Bearman, Dawson et al., 2020).
Authenticity	Authentic assessment is an "umbrella term for several important pedagogical strategies that seek to immerse learners in environments where they can gain highly practical, lifelong learning skills" (Adams Becker et al., 2018, p. 24).
Engagement	Educators are enabled to design creative and diverse assessment tasks that appropriately challenge and enable students to achieve the desired learning objectives.
Equity and fairness	The content, format and conduct of assessment are designed to ensure that no individuals or groups of students are unfairly advantaged or disadvantaged.
Feedback	Feedback is a participatory process in which learners make sense of information about their performance and use it to enhance the quality of their work or learning strategies.
Inclusion	"Assessment that recognises diversity in student learning, and endeavours to ensure that no student is discriminated against by virtue of features other than their ability to meet appropriate standards" (Tai et al., 2023, p. 484).
Support (pedagogical)	Support where the primary focus and emphasis was on pedagogical purpose, enacted through digital assessment experts working with educators in (re)designing assessment tasks using the new DAP.

Table 2
Name and description of contextual factors considered in the evaluation

Contextual factor	Description
Scalability and resourcing	The broad category of workload improvements through ease of use, automation and administrative efficiencies achieved through the DAP.
Support (other)	Support available, including technical, leadership, administrative, procedural support for staff and guidance for students that impact the pedagogical outcomes of the DAP.
Workload	Assessment is achievable for teachers (and students) with reasonable effort within the context of a full-time workload as described in institutional policy.

Method

A mixed methods approach was undertaken for data collection to provide a general understanding of the DAP adoption and use to transform assessment practices. Participants included educators, learning designers, support staff and students who had used the DAP. Data collection involved a staff online survey, interviews and case studies, a student focus group, a staff-student partnership project and formal student evaluations of courses.

Staff data collection and analysis

The research team, whose members were assessment experts, digital assessment managers and designers, and academics experienced in assessment research, designed the online survey. The survey was informed by the study's list of pedagogical criteria and the relevant contextual factors identified from literature and the university environment. After we received human ethics approval from the university, the survey was sent to 322 staff members who used the DAP for assessment. Consent to participate was seen as completing the survey online. A total of 63 course staff, including course coordinators, learning designers and tutors, responded between 10 June and 28 August 2023, equating to a 20% response rate. Descriptive statistical analysis was employed to analyse the 13 quantitative survey questions. The survey included three open-ended questions, the responses to which were thematically analysed using a coding structure related to the pedagogical criteria and contextual influences in NVivo software. Participants could opt for follow-up Zoom interviews, resulting in six interviews with nine participants. Selected case studies highlighted diverse examples of DAP usage across five individual courses and one school-based implementation.

Student data collection and analysis

Additionally, students who had used the DAP in a course in Semester 2, 2022, or Semester 1, 2023, were invited to a 60-minute online focus group in August 2023, resulting in 10 students participating. Notes were taken during the interviews and the focus group, and the sessions were audio- or video-recorded. Participants reviewed and approved the transcriptions before analysis. Interview participants were given the opportunity to edit their responses. Two additional data sources were readily available: an existing staff-student partnership (SSP) project and the routine student evaluation of all university courses from Semester 2, 2020, to Semester 1, 2023. The SSP project partnered with students to gather insights into student preferences regarding assessment questions, such as how they were presented and scaffolded, what elements were critical or "'nice to have" and how to incorporate better feedback. Despite the small student sample (n = 11), the students in the SSP project provided valuable insights to help staff use the design features of the DAP effectively for student engagement and learning.

Data were collected from the formal student course evaluations between Semester 2, 2020, and Semester 1, 2023, to glean further understanding of students' experience and perspectives of using the DAP in their courses. Analysis by the university's evaluation team focused on course experience, usefulness of tools or

materials for learning, clarity of assessment requirements, overall course rating, feedback on strengths and areas for improvement.

Results

This section reports on the staff responses from the survey, interviews and case studies and then the students' responses from the focus group, staff-student partnership project and formal evaluation of courses.

Staff perspectives

Analysis of the data revealed five main areas of reporting which align to the pedagogical and contextual factors in Tables 1 and 2: opportunities for improved assessment design, marking and feedback; improvements in assessment practices; benefits of using the DAP in one's teaching practice; matching initial reasons for using the DAP with outcomes; and pedagogical and technical support.

Opportunities to improve assessment design, marking and feedback

Educators were attracted to using the DAP for numerous reasons, whether it was part of an overall course redesign, wanting to change the type of assessment used for pedagogical reasons (Table 1) or addressing a logistical or administrative challenge (Table 2). Table 3 below provides more detail of why staff participants initially used the DAP as recorded in the survey. Respondents were given the option to select all choices that apply.

Table 3
"Why did you first start using the DAP? (select all that apply)"

Selected response	No. of responses	% of respondents
Explore a new technology	30	48%
Innovate assessment	26	41%
Create auto-marked questions to reduce workload	19	30%
Create assessment not possible on paper or the LMS[a]	14	22%
Solve an assessment administration problem	14	22%
Enhance student engagement	12	19%
Make progress towards an assessment redesign goal	10	16%
The support offered during the pilot semester	10	16%
Solve an assessment problem	9	14%
It was recommended to me	7	11%

[a] LMS = learning management system.

Survey responses indicate the diversity of how the DAP was used across the institution as shown in Table 4.

Table 4
"What forms of assessment did you use with the DAP? (select all that apply)"

Selected response	No. of responses	% of respondents
Non-invigilated exam	21	33%
Revision task or quiz	18	28%
On-campus bring-your-own-laptop digital exam	15	24%
Tutorial or problem set	11	17%
Paper or report	7	11%
Broken down step-by-step assignment	6	9%
Single or multiple case study assessment	6	9%
Laboratory practical	5	8%

Similarly, the interviews and subsequent case studies exemplify the diversity of how the DAP was used across different disciplines, with its technological affordances allowing educators to build pedagogically rich assessment tasks and significantly increase the flexibility of assessment design. The teaching staff in one case found that the validity of the new assessment design enabled them to distinguish between students who knew the right answers and those who were guessing. Another sought to focus on challenging students' higher order thinking skills through real-world genres that would be meaningful for the students. Staff reported a belief that the DAP enhanced academic integrity by allowing randomisation within sections (pools) of questions, and more easily being able to design authentic, "non-googleable" questions and tasks. Allocating specific questions to mark was also nominated as a key academic integrity benefit, as marking all of one question made it easier to detect where students had colluded.

Interviewees reported that the marking process was significantly easier, even when dealing with large numbers of markers or established marking routines. Digital marking also addressed the logistical challenge of hardcopy papers going missing. Additionally, survey results highlighted that teaching staff found it easier than reading handwritten responses. Three interviewees reported that auto-marking saved time and was cost-effective. Further, three survey respondents commented on how the DAP provides timely feedback to students through the automated grading and pre-written feedback features.

Improvements in assessment practices

The survey respondents were asked about improvements in assessment practices they had seen from using the DAP. It is interesting to note that 44% of respondents perceived an improvement in the learning experience of students, followed by a significant contextual factor in the minds of respondents, that is, maintaining workload sustainability for staff. Although recording a lesser number of responses, pedagogical improvements included "assessment authenticity" and the "feedback process" and then "alignment and validity".

When survey respondents were specifically asked about the main benefit for student learning, the connection between usability and pedagogical outcomes became more apparent. The most reported benefit was the system's modern interface that students found easy to use and provided them with authentic and personalised learning experiences, through interactive and engaging tasks.

The interviewees reported that feedback received from students was very positive around their experience with assessments in the DAP. Instructions were clear and students knew how to address the tasks. Although tasks may have challenging content, the layout and ability to break tasks down into steps assisted students with thinking through their answers. In their view, the DAP also enabled learning to be scaffolded, making students feel more confident in approaching their assessment tasks.

Benefits of using the DAP in one's teaching practice

Survey respondents were asked to identify the main benefit of using the DAP for their assessment practices. Collectively, the pedagogical criteria of engagement, authenticity, academic integrity and feedback (n = 20) ranked as benefits. However, the largest number of single-category responses identified resourcing and scalability (n = 12), which are contextual factors, as the main benefit. There were two responses that reflected little or no benefit. Table 5 presents quotations that illustrate insights in each category.

Table 5
"What do you consider to be the main benefit of using the DAP in your teaching practice?"

Category	No. of responses	Quotation examples
Resourcing & scalability	12	"Administrative ease, specifically distribution of exams to external markers, recording of results. It is much better with respect to reading student responses, and I find I can mark more exams efficiently and effectively with it." "The ability to provide feedback for students in a large course." "Auto marking."
Engagement	10	"Nontraditional question formats, upload sketches, tables and the use of video examples." "Flexible assessment formats: the DAP supports a wide range of assessment formats, such as multiple-choice questions, essays, and more. This flexibility allows educators to create diverse and engaging assessments that align with their teaching objectives."
Authenticity	5	"I can ask more dynamic and interactive questions than was possible on paper or with other platforms like the LMS." "I think its advantage is it can support multiple different assessment designs and types which allows more scope in designing authentic and engaging assessment."
Academic integrity	3	"Other than a hand-written exam that is closed book and closely supervised, it is the only way to prevent cheating and that is only so for the DAP where the exam is sat on a computer at the university rather than online (since I have experienced cheating using a phone by an offshore student who completed the exam on a computer)."
Feedback	2	"The DAP enables detailed feedback and analytics, allowing educators to provide personalised feedback to students and gain insights into their performance. This feedback can facilitate targeted interventions and help students improve their understanding of the subject matter." "The DAP allows me to deliver practice quizzes prior to the in-class exam so students can get familiar with the DAP and style of questions, while checking their own learning progress (formative); for the actual in-class exam."
No benefit	1	"There was no benefit. It was a practical experiment that I will not use again. I felt the DAP to be highly limiting in the way one had to construct the assessment."

Matching initial reasons for using the DAP with outcomes

Survey respondents evaluated how worthwhile their assessment enhancement work with the DAP was, considering their reasons for using the platform. The respondents (n = 48) were asked to record their choice using a scale of 1 (*not at all*) to 5 (*extremely worthwhile*). A total of 26 respondents indicated their assessment enhancement work was either very or extremely worthwhile. Another 16 respondents considered it to be moderately or slightly worthwhile. Six respondents indicated using the DAP was not at all worthwhile.

Although the number of respondents to this question was less than the total number of survey participants, 73% considered the assessment improvement work to be extremely, very or moderately worthwhile, with 54% judging their assessment enhancement work with the DAP to be extremely or very worthwhile.

In the second part of this question, respondents were asked to explain their rating. Positive comments focused on the pedagogical features of academic integrity, engagement, feedback and support. Staff workload was the key contextual factor, especially in the initial stages of learning the platform. A few respondents found the significant time investment required to learn the DAP too high for the resulting benefits, opting to continue using the LMS or other existing systems. Other comments highlighted instances where the experience of using the DAP did not meet expectations, such as challenges with multiple markers, feedback limitations and technological issues for students using certain operating systems.

Pedagogical and technological support

Respondents learnt to use the DAP and develop digital assessment on the platform in a variety of ways. The survey responses showed that learning how to effectively use the new platform and develop assessment using it required collaboration. Respondents reported collaborating with the central digital assessment team, school or faculty-based learning designers, other academic colleagues, and/or with school or faculty professional staff. In addition to co-designing and building assessment with central support staff, respondents also mentioned several resources offered by the central teaching and learning unit, including hosted workshops and events, using online guides maintained by this unit, and eLearning systems and support teams. Others had their assessments built for them by the central support team.

Support also emerged as a consistent theme across the interviews, especially the acknowledgement of the digital assessment team in the central teaching and learning centre for their support during the initial learning curve. Ongoing support was found to be valuable in different ways, including colleagues, faculty learning designers and self-directed learning. At a school-level implementation, school leadership support was seen as a vital factor. Staff acknowledged concerns around an initial increase in workload to learn how to effectively use the DAP, but many considered that the support provided help offset the possible costs due to contextual factors identified in Table 2.

Student perspectives

Similar headings to those used in the Staff perspectives section are used to report student experiences with the DAP, highlighting pedagogical and contextual factors from three data sources, the SSP, focus group and course evaluations.

Opportunities to improve assessment design, experience and feedback

The staff-student partnership sought to understand student perspectives on using the DAP, to support staff in utilising the platform in ways that best help students meaningfully engage and learn through digital assessment and feedback. Students were asked questions about their assessment preferences, including presentation, key elements and feedback improvements. For example, they were asked if it would be helpful to add a “Before you submit checklist” at the end of an assessment task prior to submitting. Students indicated the top benefits of such a checklist would be that “it provides confidence you haven’t forgotten anything, especially if you are rushing and it provides a critical summary of the requirements” and “it’s easier to read than a text heavy task sheet”. Further, all participants stated that it was either extremely or very important not to have to scroll to see information required to answer questions. Several indicated a preference for keyboard entry answers where possible, rather than having to drag to categorise or place information or answers into spaces or boxes. Nearly the same number of respondents said they would prefer to keep using scratch paper in digital on-campus examinations.

These SSP students were asked whether they saw benefit in adding a feedback form to the end of an assessment task prior to submitting an assignment. Most students indicated that they would find it motivating if they received quality feedback that addressed areas in which they requested feedback, and 54% said it would help with feeling seen by their lecturer, but they were surprisingly non-committal about whether they thought feedback forms should be a standard inclusion in assignments or whether they would actually fill in the form, with the most common response to both questions being “maybe”. Students overwhelmingly reported that they had never received any feedback in the DAP other

than whether they got a question correct or incorrect, which shows staff may not using the multiple ways of providing feedback to students in the DAP platform.

Improvements in assessment experience

Students in the focus group discussed the positive features of online assessments, particularly in examination circumstances, and tended to focus on the technological affordances of the DAP. Students agreed they found typing more convenient and efficient, enabling them to edit and format their responses easily, as explained by one participant:

> It's more stressful in a written exam, for example, because one thing I always think about in a written exam is, can the teacher like even understand my handwriting.

Focus group students also valued the interface, organisation, and easy navigation opportunities of the DAP:

> I definitely think the DAP makes it much easier to just upload. Submit. And then that's it. You're done.
> I prefer the whole interface in the DAP and it's easier to track which questions are where.

One student mentioned that using the DAP for practical assessments was beneficial as the assessment design and implementation enabled students to complete the assessments even if they could not attend the physical practical sessions:

> It would be open on the day of the prac, and then during the prac, which is a 3-hour class you could work through the questions, and the teachers help you … and there were real life samples to look at as well. And you could do it at home if you wanted to and then you'd have a whole week to submit it. It was just really convenient.

Technological support

Students in the focus group also recognised the challenges associated with digital assessment, including concerns around what the DAP might do to their laptop and the technological issues they might experience, as explained by two students:

> To be totally honest when we first were told we were going to use it for our mid Sem [invigilated exam], I wasn't super happy with the idea of like the software itself just because it has the ability to access your microphone and your camera and lockdown your computer.

> And even though I've had technological difficulties (the red screen) there's always people to help us, and we can also have extra time if we really experience this kind of difficulties.

Student evaluation of using the DAP in their courses

Five questions, from the standard 10 course evaluation set, were seen as most relevant to provide insight about the DAP's influence on student learning as explained in Table 6. There were 7,279 responses from 187 unique courses that had used the DAP. Descriptive statistical analysis of these courses mainly received a score of 4 or 5 out of 5 in the quantitative questions – 4 (mean 4.24), 5 (mean 4.15) and 8 (mean 4.16) – indicating that most students agreed that course experiences, tools or materials were made clear to them. This pattern was also observed for all the university courses in the same period. In summary, the course quantitative analysis of these questions shows that generally courses that used the DAP had similar levels of student satisfaction as courses that did not participate in the project.

Table 6
Formal student evaluation of course questions used for DAP

Question no.	Question description	Response options
4	Course experience, tools or materials were useful for my learning	A scale of 1–5, with 5 being *strongly agree* and 1 being *strongly disagree*
5	Assessment requirements were made clear to me	A scale of 1–5, with 5 being *strongly agree* and 1 being *strongly disagree*
8	Overall, how would you rate this course?	A scale of 1–5, with 5 being *strongly agree* and 1 being *strongly disagree*
9	What were the best aspects of this course?	Qualitative responses
10	What improvements would you suggest?	Qualitative responses

There were 4,489 responses to Question 9, which asked the students to list the best aspects of the course and 4,333 responses to Question 10, which asked about improvements. The qualitative analysis involved searching for the term *Name of the DAP* in the course comments and attaching a positive or negative sentiment to those comments. During the 6 semesters in question, only 130 comments mentioned the DAP platform. A total of 31 comments were found for Questions 9 and 99 for Question 10, respectively. In Question 9, of the 31 comments, 30 had a positive sentiment with students commenting on ease of navigation, clear layout and the learning benefit they felt they achieved through regular assessment tasks, especially for use in practicals. Several comments indicated a preference for the DAP over the LMS. Conversely, in response for suggested improvements, 85 out of 99 comments had a negative sentiment. Common experiences included technological difficulties, heightened stress levels during examinations using the DAP, layout issues and access problems for feedback or assessment reviews. Students also noted inconsistencies between the information in the DAP and other course materials, expressing reservations about lockdown browsers for open-book examinations.

Whilst students reported mixed experiences using the DAP, the differences are likely to be attributed to how the implementation was perceived at an individual course level. It should also be noted that stress related to technological difficulties (or the anticipation of them) is a factor in all digital examinations at the university not just those using the DAP.

Discussion and future directions

Staff and student perspectives on using the DAP for assessment

The study sought to gain insights into the pedagogical benefits and challenges of using the DAP across a large metropolitan university through the perspective of both staff and students. Results highlight strong pedagogical reasons for continuing the digital assessment initiative but also some constraints and opportunities for improvement. Staff were attracted to using the DAP for pedagogical and non-pedagogical reasons, including enhancing exam assessment as well as a variety of other assessment types. They enjoyed the possibility of developing flexible, inclusive, authentic and engaging tasks, within a more secure assessment environment. Notably, staff commended the platform's formative and dialogic feedback features, though student feedback from the SSP project indicated room for improvement in this area. Staff focused on pedagogical benefits for student learning (cf. Mimirinis, 2018) within workload and resource constraints (cf. St-Onge et al., 2022), while students were more concerned with platform usability, technological issues and the completion of their assessments.

Although creativity in achieving pedagogical outcomes is important, educators must navigate current contextual constraints within higher education. Balancing the creativity enabled through the DAP with workload implications is essential (cf. Bennett et al., 2017). Educators need to consider how much creativity in the design is necessary to achieve the desired learning outcomes. The SSP project provided insights into student perspectives on effective ways to use the DAP to garner a response, rather than incorporating elaborate features. Perhaps a simpler design requiring less workload, and clever use of functionality options such as re-use and auto-assembly of content, auto-marking and/or using feedback

options may be more sustainable. Achieving these efficiencies, however, usually requires an initial investment of time, which can be difficult for staff with time and budgetary constraints. Participants in this study commented on the necessity of adequate support (cf. Bennett et al., 2017) to assist them in the initial learning curve in using the platform but of particular importance, to (re)design assessment that is pedagogically effective. This suggests the need for expert design support staff with comprehensive knowledge of the system and literature-informed best practice to work alongside the teaching staff.

Students emphasised the ease of using the DAP, aiming to reduce the stress of examinations by focusing on usability, supporting elements of the TAM user experience models (Davis, 1986). Although they appreciated the digital format, concerns about accessing the Internet, disappointing computer performance and navigating a digital environment during high-stakes assessments remained. Familiarisation opportunities were highlighted as beneficial for students to make full use of the digital platform not only for completing assessment, but also for accessing feedback, revising and setting new learning goals based on past assessment performance. Although this is not unique to this DAP, these sessions can assist students to be more prepared and confident in their assessment performance.

The impact of contextual factors on pedagogical outcomes

We reviewed literature to find frameworks for a pedagogical evaluation, selecting the Huber et al. (2023) framework, which highlighted contextual factors, such as scalability, resource availability and institutional policies that affect pedagogical outcomes in digital assessment. Yet the framework was not entirely suited to the DAP initiative, so we used it to develop key factors of assessment and context to guide data collection and analysis, considering scalability and resourcing, support (other than pedagogical) and workload. The results show that a purely pedagogical evaluation is challenging because contextual, technical and administrative factors also influence outcomes; identifying clear pedagogical outcomes proved difficult due to the overlap of usability, context and pedagogy, further complicated by contrasting staff and student perspectives. Taking a multifaceted and multi-sourced approach (Chalmers & Hunt, 2016) combined with the guiding key factors nevertheless proved fruitful for working with these challenges and providing a rich picture of how the DAP was experienced.

Impacting contextual factors can be significant at an institutional decision-making level. Ensuring the DAP's longevity is crucial considering the significant effort educators and students put into DAP. Competing interests in a resource-limited educational environment can lead to disruption in the introduction or continuation of new digital systems. It is advisable to avoid frequent system changes, as this allows staff to capitalise on their time invested in assessment design and implementation over an extended period. Institutions cannot make a habit of asking educators to try new digital platforms without the guarantee that they will still be able to use it in the future. The financial and support investment should be seen for the long term (at least 5 years) in digital systems, such as this DAP that requires an initial steep learning curve but produces quality assessment outcomes and ongoing efficiencies. As a result, evaluations of this kind are becoming a necessary part of any teaching and learning initiative (Chalmers & Gardiner, 2015), providing evidence of the pedagogical value of the system.

The limitation of this work is its exploratory and practical nature within only one institution. However, it advances the research in this field by providing an approach to the design of a pedagogical evaluation, potential available data collection opportunities and the result that pedagogical factors will always be influenced and intermingled with technological and contextual factors (cf. Mishra & Koehler, 2006). In the context of digital assessment, models and frameworks that enable scrutiny of assessment practices, resources and processes relevant to centrally supported initiatives are relatively scarce, and those that address aspects of digital assessment beyond functionality, usability and user experience are even rarer. Research is needed to build and validate a framework that meets the needs of pedagogical evaluations and their relationships with other contextual factors. In a more practical vein, examples of other pedagogical evaluation approaches at an institutional level would assist practitioners who are asked to undertake similar work.

Conclusion

This paper presents a pedagogical evaluation approach for a DAP at a large Australian metropolitan university. Existing user experience models evaluate the effectiveness of technologically enhanced initiatives. Measuring the pedagogical impact of digital assessment initiatives has not been so straightforward. The collected data provides comprehensive insights into DAP users' pedagogical perspectives from both staff and student viewpoints, aligning with project timelines and available resources. Results show there are robust pedagogical justifications for continued DAP usage, as it enhances student experiences through improved assessment design and interface usability. However, there are also some constraints and opportunities for improvement. The insights generated through this evaluation also highlight the importance of considering technical and pedagogical factors in relation to each other and to contextual factors. In a quality assurance climate where the demand for this type of institution-wide evaluation is increasing, this study contributes an approach that addresses some of the complexities and challenges of evaluating digital assessment initiatives in higher education.

Author contributions

Author 1: Conceptualisation, Investigation, Formal analysis, Writing – original draft, Writing – review and editing, Response to reviewers; Copyright review **Author 2**: Investigation, Formal analysis, Writing – original draft, Writing – review and editing; Response to reviewers; **Author 3**: Investigation, Formal analysis, Writing – original draft, Writing – review and editing; **Author 4**: Conceptualisation, Investigation, Formal analysis, Writing – original draft, Writing – review and editing; **Author 5:** Conceptualisation**,** Investigation, Formal analysis, Writing – original draft, Writing – review and editing; **Author 6:** Conceptualisation, Investigation, Formal analysis, Writing – original draft; **Author 7:** Conceptualisation, Investigation, Formal analysis, Writing – original draft, Writing – review and editing.

Acknowledgements

The authors would like to thank all the staff and student participants in this project.

References

Adams Becker, S., Brown, M., Dahlstrom, E., Davis, A., DePaul, K., Diaz, V., & Pomerantz, J. (2018). *NMC Horizon Report: 2018 Higher Education Edition*. EDUCAUSE. https://library.educause.edu/~/media/files/library/2018/8/2018horizonreport.pdf

Ahmed, M. R., & Sidiq, M. A. (2023). Evaluating online assessment strategies: A systematic review of reliability and validity in e-learning environments. *North American Academic Research, 6*(12), 1–18. https://doi.org/10.5281/zenodo.10407361

Appiah, M., & van Tonder, F. (2018). E-assessment in higher education: A review. *International Journal of Business Management and Economic Research, 9*(6), 1454–1460. https://www.ijbmer.com/docs/volumes/vol9issue6/ijbmer2018090601.pdf

Bearman, M., Boud, D., & Ajjawi, R. (2020). New directions for assessment in a digital world. In M.Bearman, P. Dawson, R. Ajjawi, J. Tai, & D. Boud (Eds.), *The enabling power of assessment: Vol. 7. Re-imagining university assessment in a digital world* (pp. 7–18). Springer.

Bearman, M., Dawson, P., O'Donnell, M., Tai, J. & Jorre de St Jorre, T. (2020). *Ensuring academic integrity and assessment security with redesigned online delivery*. Deakin University. https://dteach.deakin.edu.au/wp-content/uploads/sites/103/2020/03/DigitalExamsAssessmentGuide1.pdf

Bennett, S., Dawson, P., Bearman, M., Molloy, E., & Boud. (2017). How technology shapes assessment design: Findings from a study of university teachers. *British Journal of Educational Technology, 48*(2), 672–682. https://doi.org/10.1111/bjet.12439

Bloxham, S., & Boyd, P. (2007). *Developing effective assessment in higher education: A practical guide*. Open University Press. https://insight.cumbria.ac.uk/id/eprint/343/1/Bloxham_DevelopingEffectiveAssessment.pdf

Boud, D., & Associates (2010). *Assessment 2020: Seven propositions for assessment reform in higher education*. Australian Learning and Teaching Council. https://www.kpu.ca/sites/default/files/Teaching%20and%20Learning/Assessment_2020_final%5B1%5D.pdf

Bovill, C., Matthews, K. E., & Hinchcliffe, T. (2021). *Student Partnerships in Assessment (SPiA)*. Advance HE. https://documents.advance-he.ac.uk/download/file/document/10140?_ga=2.151925050.2076618188.1728360152-1141919176.1728360152

Celik, V., & Yesilyurt, E. (2013). Attitudes to technology, perceived computer self-efficacy and computer anxiety as predictors of computer supported education. *Computers & Education*, *60*(1), 148–158. https://doi.org/10.1016/j.compedu.2012.06.008

Chalmers, D., & Gardiner, D. (2015). An evaluation framework for identifying the effectiveness and impact of academic teacher development programmes. *Studies in Educational Evaluation, 46*, 81–91. https://doi.org/10.1016/j.stueduc.2015.02.002

Chalmers D., & Hunt, L. (2016). Evaluation of teaching. *HERSDA Review of Higher Education, 3*, 25–55. http://herdsa.org.au/herdsa-review-higher-education-vol-3/25-55

Davis, F.D. (1986). *A technology acceptance model for empirically testing new end-user information systems: Theory and results.* [Doctoral dissertation Massachusetts Institute of Technology]. DSpace@MIT. https://dspace.mit.edu/handle/1721.1/15192

Dawson, P. (2020). *Defending assessment security in a digital world: preventing e-cheating and supporting academic integrity in higher education* (1st ed.). Routledge. https://doi.org/10.4324/9780429324178

Fluck, A. E. (2019). An international review of eExam technologies and impact. *Computers & Education*, *132*, 1–15. https://doi.org/10.1016/j.compedu.2018.12.008

Gerritsen-van Leeuwenkamp, K. J., Joosten-ten Brinke, D., & Kester, L. (2017, December). Assessment quality in tertiary education: An integrative literature review. *Studies in Educational Evaluation, 55*, 94–116. https://doi.org/10.1016/j.stueduc.2017.08.001

Gneil, H. (2023, June 7). *ChatGPT Webinar #3: What have we learnt?* [Video]. YouTube. https://youtu.be/1V0f3pw2B_0?t=686

Guskey, R. T. (2000). *Evaluating professional development.* Corwin Press Inc.

Henderson, M., Selwyn, N., & Aston, R. (2017). What works and why? Student perceptions of 'useful' digital technology in university teaching and learning. *Studies in Higher Education*, *42*(8), 1567–1579. https://doi.org/10.1080/03075079.2015.1007946

Huber, E., Harris, L., Wright, S., Raduescu, C., White, A., Raduescu, C., Zeivots, S., Cram, A., & Brodzeli, A. (2023). Towards a framework for designing and evaluating online assessments in business education. *Assessment & Evaluation in Higher Education*, 49(1), 102–116. https://doi.org/10.1080/02602938.2023.2183487

International Center for Academic Integrity. (2021). *The fundamental values of academic integrity* (3rd ed.). https://academicintegrity.org/images/pdfs/20019_ICAI-Fundamental-Values_R12.pdf

Kemp, A., Palmer, E., & Strelan, P. (2019). A taxonomy of factors affecting attitudes towards educational technologies for use with technology acceptance models. *British Journal of Educational Technology, 50*(5), 2394–2413. https://doi.org/10.1111/bjet.12833

Mercader, C., & Gairín, J. (2020). University teachers' perception of barriers to the use of digital technologies: the importance of the academic discipline. *International Journal of Educational Technology in Higher Education*, *17*, Article 4. https://doi.org/10.1186/s41239-020-0182-x

Mimirinis, M. (2019). Qualitative differences in academics' conceptions of e-assessment. *Assessment & Evaluation in Higher Education, 44*(2), 233–248. https://doi.org/10.1080/02602938.2018.1493087

Mishra, P., & Koehler, M.J. (2006). Technological pedagogical content knowledge: A framework for teaching knowledge. *Teachers College Record, 108*(6), 1017–1054. https://doi.org/10.1111/j.1467-9620.2006.00684.x

Ryan, T., Henderson, M., Ryan, K., & Kennedy, G. (2021). Designing learner-centred text-based feedback: a rapid review and qualitative synthesis. *Assessment & Evaluation in Higher Education*, *46*(6), 894–912. https://doi.org/10.1080/02602938.2020.1828819

Sadler, D. R. (2005). Interpretations of criteria-based assessment and grading in higher education. *Assessment & Evaluation in Higher Education*, *3*(2), 175–194. https://doi.org/10.1080/0260293042000264262

Sadler, D. R. (2017). Assuring academic achievement standards: from moderation to calibration. In V. Klenowski (Ed.), *International teacher judgement practices* (1st ed., pp. 15–29). Routledge.

Schellekens, L. H., van der Schaaf, M. F., van der Vleuten, C. P. M., Prins, F. J., Wools, S., & Bok, H. G. J. (2023). Developing a digital application for quality assurance of assessment programmes in higher education. *Quality Assurance in Education, 31*(2). 346–366. https://doi.org/10.1108/QAE-03-2022-0066

St-Onge, C., Ouellet, K., Sawsen, L., 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

Sweeney, T., West, D., Groessler, A., Haynie, A., Higgs, B., Macaulay, J., Mercer-Mapstone, L., & Yeo, M. (2017). Where's the transformation? Unlocking the potential of technology-enhanced assessment. *Teaching and Learning Inquiry, 5*(1), 41–64. https://doi.org/10.20343/teachlearninqu.5.1.5

Tai, J., Ajjawi, R., Bearman, M., Boud, D. Dawson, P., & Jorre de St Jorre, T. (2023). Assessment for inclusion: Rethinking contemporary strategies in assessment design. *Higher Education Research & Development. 42*(2), 483–497. https://doi.org/10.1080/07294360.2022.2057451

Teo, T. (2019). Students and teachers' intention to use technology: Assessing their measurement equivalence and structural invariance. *Journal of Educational Computing Research*, *57*(1), 201–225. https://doi.org/10.1177/0735633117749430

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://doi.org/10.1002/berj.3215

Venkatesh, V., & Bala, H. (2008). Technology Acceptance Model 3 and a research agenda on interventions. *Decision Sciences*, *39*(2), 273–315. https://doi.org/10.1111/j.1540-5915.2008.00192.x

Venkatesh, V., & Davis, F.D. (2000). A theoretical extension of the technology acceptance model: Four longitudinal field studies. *Management Science, 46*(2), 186–204. https://doi.org/10.1287/mnsc.46.2.186.11926

Wang, X., Su, Y., Cheung, S., Wong, E., & Kwong, T. (2013). An exploration of Biggs' constructive alignment in course design and its impact on students' learning approaches. *Assessment & Evaluation in Higher Education*, *38*(4), 477–491. https://doi.org/10.1080/02602938.2012.658018

Winstone, N., & Carless, D. (2020). *Designing effective feedback processes in higher education: A learning-focused approach*. Routledge.

Yan, Z., Wang, X., Boud, D., & Lao, H. (2023). The effect of self-assessment on academic performance and the role of explicitness: a meta-analysis. *Assessment & Evaluation in Higher Education*, *48*(1), 1–15. https://doi.org/10.1080/02602938.2021.2012644

Corresponding author: Christine Slade, c.slade@uq.edu.au

Please cite as: Slade, C., Mahon, K., Benson, K., Lynagh, J., McGrath, D., Sheppard, K., & Ahsan, Q. (2024). A pedagogical evaluation of an institution's digital assessment platform (DAP): Integrating pedagogical, technical and contextual factors. *Australasian Journal of Educational Technology*, *40*(4), 90–104. https://doi.org/10.14742/ajet.9448