Impact of Formative E-Assessment on Attendance

Mario Orsi
University of the West of England, Bristol

Leonardo Juliano
University of the West of England, Bristol

E-assessment is an established and growing approach that can offer several advantages compared to traditional pen and paper assessments, such as instant feedback, online/remote interaction, and automated personalization. In this study, we investigate the impact of introducing formative e-assessment in a compulsory science module of an undergraduate degree programme. In particular, we compare two paired cohorts, from consecutive academic years, before and after replacing pen and paper formative assessment activities with digital alternatives based on the Numbas e-assessment platform. We find that introducing e-assessment correlates with a marked and statistically significant increase in student attendance. Module evaluation comments from students indicate an overall view of e-assessment as beneficial and engaging.

Keywords: e-assessment, attendance, formative assessment, online learning, personalization, paired t-test

INTRODUCTION

E-assessment (also known as online, digital, computer-based or web-based assessment), generally defined as formative/summative assessment relying on information technology, offers significant potential advantages over traditional pen and paper approaches, in particular with regards to authenticity, flexibility and automation. Also considering the increasing reliance on digital devices and platforms in society at large, it is unsurprising that e-assessment is becoming more and more prevalent and desirable. Moreover, this tendency has accelerated greatly as a consequence of the Covid-19 lockdowns and related need for remote online assessment (Watermeyer et al. 2020; Rapanta et al. 2020).

A recent systematic review on the subject highlights a general student preference for e-assessments over traditional paper exams; in particular, most students perceive online learning as more supportive with regards to wellbeing, personal lives, and learning performance (Butler-Henderson and Crawford 2020). Notably, students appreciate that e-assessment is more authentic (Williams and Wong 2009) and more consistent with real-world work settings (Matthiasdottir and Arnalds 2016). It is indeed worth stressing how traditional exams, which require writing by hand, on paper, in invigilated rooms with no Internet access, represent a stridently different experience compared to today's graduates' typical work environment, which overwhelmingly includes computers, access to the Internet, and often specialized application software.

Previous research shows that e-assessment is altogether advantageous compared to pen and paper also from the teachers' perspective (Pagram et al. 2018; Rolim and Isaias 2019; Mimirinis 2019; Casey and

Crowley 2019), despite the presence of some challenges (Kuikka, Kitola, and Laakso 2014; Gillett-Swan 2017).

Regarding student performance, e-assessment was found to have a significant positive effect in two separate large-scale studies of business mathematics courses (Angus and Watson 2009; Weir, Gwynllyw, and Henderson 2021), although no significant effect was detected for a (smaller) study of a chemistry course (Prisacari and Danielson 2017).

In general, it is important to note that any study on potential impacts of e-assessment on students' learning, behaviour or performance, as well as on students/teachers perceptions, is inevitably biased by specific study features, such as the e-assessment tools used, assessment details (e.g., basic 'static' multiple choice questions vs. elaborate multistep problems with randomized 'dynamic' parameters), the student sample, the discipline, etc. While this makes it difficult to establish general principles, individual studies are of course valuable in their specific context as well as cumulatively, especially with regards to any common finding.

In this article, we investigate the effect of introducing formative e-assessment in relation to attendance. We consider two cohorts of students from consecutive academic years; the first one is exposed to traditional pen and paper formative assessment activities, whereas in the second one such activities are replaced by corresponding e-assessment alternatives. Other relevant variables are kept the same, including module content, delivery, teaching staff, and session timetables (relative to the specific academic year). Student attendance has been repeatedly reported to strongly correlate with performance (Crede, Roch, and Kieszczynka 2010; Andrietti 2014; Dey 2018; Tetteh 2018; Edwards and Clinton 2019; Buchele 2021), although studies exist which do not find a significant relation (Azab et al. 2016; Kauffman et al. 2018). Attendance is also found to correlate with student satisfaction (Kuh et al. 2011), and is regarded as an important quality indicator (Coates 2005). For these reasons, unsurprisingly student attendance is universally promoted and monitored (to various degrees) by higher education institutions. Previous studies found evidence of increased engagement following the introduction of low-stakes continuous e-assessments (Carroll et al. 2017; Holmes 2015, 2018). However, in these studies it is difficult to disentangle the likely extra incentive of introducing a summative element combined with the e-assessments themselves. Our study is different in that the e-assessment we use as comparative factor is formative (i.e., it does not contribute a mark), so arguably we can more clearly isolate any specific effect that introducing e-assessment has on attendance.

HYPOTHESIS

We propose the hypothesis that the replacement of pen and paper formative assessment with e-assessment correlates with changes in student attendance. We treat this as a two-sided hypothesis, in that we do not make an a priori assumption on the direction of any significant change (i.e., an increase vs. a decrease in attendance).

METHODS

Module Details and Context
The module studied is a compulsory second year (level 5) undergraduate course on introductory statistics, part of the 3-year Healthcare Science BSc degree apprenticeship programme at the University of the West of England (UWE Bristol), UK. A specific feature of this programme is that students are employed in the UK National Health Service (who provides the funding) and are at an early (apprentice) stage. We use data from two cohorts of students in subsequent academic years, respectively 2018-2019 and 2019-2020. The module focuses primarily on functioning (rather than declarative) knowledge; as such, students engage extensively in problem solving (e.g., applications of statistical testing). In 2018-2019, no e-assessment was used; for each session, students were given problem solving question sheets (for practice and formative assessment purposes). In 2019-2020, the question sheets from the previous year were replaced by digital versions within the chosen e-assessment platform (more details on this are given below).

Cohorts

The number of students registered on the module was 21 in the 2018-2019 academic year (when no e-assessment was used) and 31 in 2019-2020 (when e-assessment was introduced). For simplicity, we henceforth refer to the 2018-2019 and 2019-2020 cohorts as the *no e-assessment* and *e-assessment* cohorts, respectively. The *no e-assessment* cohort consists of 57.1% females and 42.9% males, while in the *e-assessment* cohort there are 51.6% females and 48.4% males; $\chi 2$ testing shows no evidence of a significant association between gender and cohort ($\chi 2 = 0.154$, $p = 0.695$).

Attendance

The module considered runs over two terms (Autumn and Spring) and is delivered with a blended approach: out of the fourteen teaching and learning sessions in total, twelve sessions take place online (six per term with a fortnightly frequency) while two sessions are delivered on campus (one at the beginning of each of the two terms). In this study we focus on the online sessions only; in these sessions attendance is not compulsory, as is conventional in higher education. We do not include campus sessions because in this case attendance is compulsory (and indeed all students were present in all sessions in both the academic years under investigation).

Online attendance data for each session were obtained by downloading attendance reports which are automatically generated by our virtual learning environment Blackboard (Blackboard Inc.). To compare the two cohorts (which have different sizes), as well as for improved clarity, the attendance for each session is quantified as a percentage (Nordmann et al. 2019) with the formula: $100 \times$ *number of attendees / cohort size*. We decided to discard the last session of the 2019-2020 cohort, as an uncontrollable external factor, the Covid-19 pandemic, caused attendance to drop to an exceptionally low rate of 32.3% (compared to an average of 63.1% for the academic year up to that point). We can confidently attribute this outlier to the Covid-19 emergency as all the students are employed in the UK National Health Service, which at the time experienced sudden severe stress due to the pandemic. In particular, the UK went into its first national lockdown on 23/3/2020 and the last session (the one which we discard) took place on 1/4/2020.

We thus consider eleven sessions for each cohort. Such sessions are 'matched' in pairs, i.e., we consider eleven pairs of observations consisting of the 2018-2019 *no e-assessment* and the 2019-2020 *e-assessment* attendance data for each session. The statistical analysis is therefore conducted with a paired t–test. The paired t-test assumes that the differences between paired observations are normally distributed; we checked this assumption using the Anderson-Darling normality test and found no evidence to reject the null hypothesis of normality ($p = 0.247$).

The Numbas E-Assessment Platform

In this study, we used the open-source e-assessment system Numbas (www.numbas.org.uk). Numbas is an established and widely-used platform; it has especially useful features for maths-based disciplines (including science, engineering and economics) but it also supports generic question types (e.g., multichoice, true/false, fill in blank) that can be used for any subject (Foster, Perfect, and Youd 2012; Graham 2020; Perfect 2015). In particular, a key feature of Numbas is the ability to automatically randomize numerical input values each time that a question is accessed by a different student (or by the same student, e.g., for repeated practice). The order and sampling of question sections can also be randomized; for example, for a 'true/false' type question, a subset of statements can be randomly sampled from a larger set. Feedback is also automated and instantaneous; notably, Numbas can automatically recognize errors 'carried forward' from previous steps and take this into account with regards to marking and feedback.

An example of the use of Numbas in this study is reported in Figure 1, which shows a comparison between a question used for formative assessment in the *no e-assessment* cohort (2018-2019) and the equivalent question implemented in Numbas and used in the *e-assessment* cohort (2019-2020).

The *no e-assessment* format (A) is static (numerical values are the same for all students), non-interactive, and feedback is asynchronous. The *e-assessment* format (B) is dynamic (numerical values are randomized each time a student accesses the question), interactive (answers are entered in the relevant boxes), and feedback is instantaneous (not shown here for clarity – see examples at www.numbas.org.uk)

FIGURE 1

COMPARISON BETWEEN FORMATIVE ASSESSMENT IN THE TWO COHORTS CONSIDERED

An independent random sample of size $n = 20$ is selected from an approximately normal population with unknown standard deviation.

a. What is the critical t-value ($t*$) for a 95% confidence level?

b. Assuming a two-tailed test, what is the p-value for a statistic $T = 2.74$?

A. Example of question for the *no e-assessment* cohort (2018-2019). The question is accessed as a standard document and all numerical values are fixed.

An independent random sample of size $n = 11$ is selected from an approximately normal population with unknown standard deviation.

a. What is the critical t-value ($t*$) for a **99%** confidence level?

[ ]

Submit part

Score: 0/5
Unanswered

b. Assuming a two-tailed test, what is the p-value for a statistic $T = 1.83$?

[ ]

Submit part

Score: 0/5
Unanswered

B. Example of question for the *e-assessment* cohort (2019-2020). The question, created with Numbas, is accessed as a web page, is interactive, and all numerical values are randomized every time the question is loaded.

RESULTS AND DISCUSSION

Effect of E-Assessment Introduction on Attendance

Attendance data comparing the *no e-assessment* (2019-2019) and *e-assessment* (20192020) cohorts are reported in Figure 2.

It can be seen that the introduction of e-assessment correlates with increased attendance for almost all (ten out of eleven) of the paired sessions considered; only one session deviates from this trend (session 6).

Descriptive statistics for the average session attendance over the academic year (eleven sessions in total) for the corresponding cohorts are reported in Table 1.

TABLE 1
DESCRIPTIVE STATISTICS FOR AVERAGE SESSION ATTENDANCE FOR THE TWO COHORTS CONSIDERED

Academic year	Statistics (as percentages)					
	Mean	Median	SD	IQR	Minimum	Maximum
2018-2019 (no e-assessment)	52.8	52.4	10.9	23.8	38.1	71.4
2019-2020 (e-assessment)	63.1	58.1	9.28	16.1	51.2	80.7

It can be seen that both mean and median attendance are higher for the *e-assessment* cohort; moreover, in this cohort attendance never falls below 50%, and reaches a maximum of 80.7%. In contrast, the *no e-assessment* cohort is characterized by a lower minimum of 38.1% as well as a lower maximum of 71.4%.

FIGURE 2
ATTENDANCE DATA: COMPARISON BETWEEN PERCENT ATTENDANCE BETWEEN THE NO E-ASSESSMENT AND E-ASSESSMENT COHORTS IN EACH PAIRED SESSION
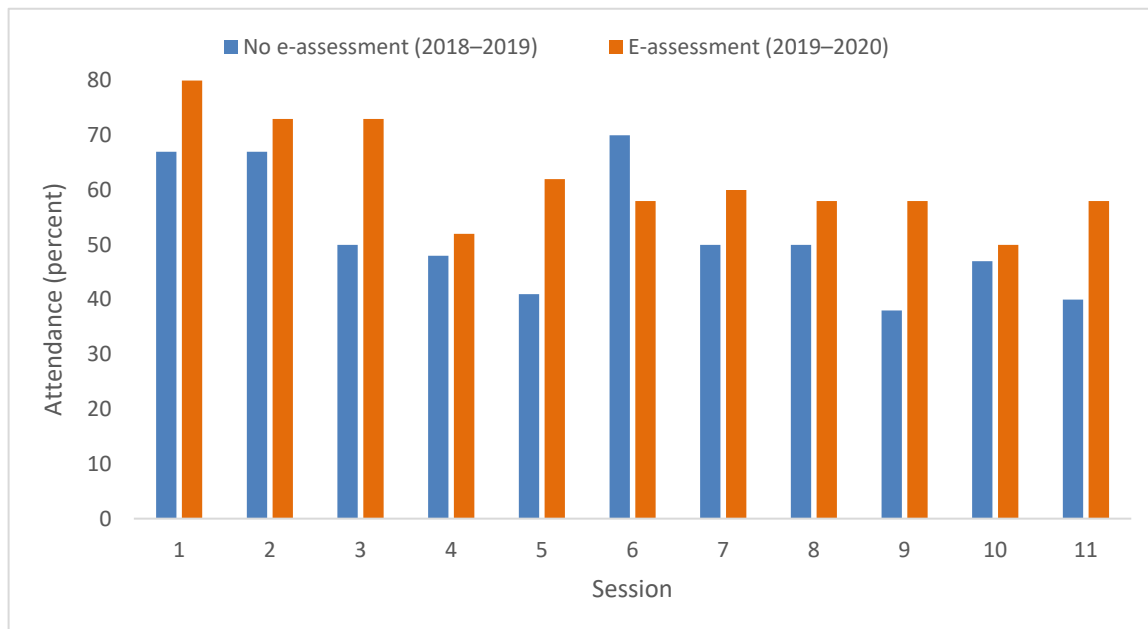
A paired t-test confirms that there is strong evidence of significantly increased average attendance in the *e-assessment* cohort with respect to the *no e-assessment* cohort (t(10) = 3.34, p = 0.007). In particular,

the mean attendance increase is 10.2% and the corresponding 95% confidence interval is between 3.4% and 17.1%. Moreover, the relative attendance increase is 19.5%.

Of course statistical significance, while necessary to accept a hypothesis, can be achieved (as long as samples are large enough) even when the effect of interest is small and of arguably no practical value. For our study, it is important to stress that the observed increase in student attendance, of approximately 10% and 20% in absolute and relative terms (respectively), is large enough to also have clear practical significance. In particular, the observed impact of introducing e-assessment corresponds to one extra student attending for every ten registered, and to one extra student attending in the *e-assessment* cohort for every five students attending in the *no e-assessment* cohort.

Evaluation From Students

To investigate, at least qualitatively, possible reasons behind the observed attendance improvement in the *e-assessment* cohort, it is useful to consider anonymous student comments submitted at the end of the module as part of the general module feedback. In particular, six students made comments related to the e-assessments, as follows:

> The online assessments worked well and the fact that you get an instant result is also beneficial. I wouldn't change anything about this assessment format.

> The online assessment is clear and simple to use. The practice questions are essential to make us put in the extra work and enable us to fully understand the subjects.

> It is extremely helpful to have the randomised questions.

> I must say the module content is excellent, especially the way we can practice randomised questions.

> The assessments were great. I found them very fair and useful, and the module was far clearer to me as a student.

> The Numbas online assessment was good. It was very useful to have exam-like questions beforehand to make the nature of the assessment clearer and more fair for all students taking it as there is less ambiguity in how to answer each question.

It is evident, at least from these statements, that students responded very positively to the introduction of e-assessment, which they appreciated especially for the instant feedback, the opportunity for extra practice through randomized questions, and the corresponding incentive to 'put in the extra work'. Overall, it appears that students found e-assessment engaging, and this is reasonably expected to promote attendance.

LIMITATIONS

A notable strength of our study is that student attendance was observed before and after the introduction of e-assessment while all other key features of the module considered were kept constant. However, it remains possible that the effect found (increased attendance upon e-assessment introduction) is due (at least partly) to unobserved and uncontrolled differences in the cohorts. Another issue is that, as mentioned earlier, the students considered are all in permanent employment (as apprentices), which is an atypical situation for undergraduate students. However, this arguably also represents a strength of our study, as it makes the sample considered rather homogeneous, i.e., all students are employed in the UK National Health Service and are all at an early (apprentice) stage; thus, they all have similar incomes and work commitments. This is unlike typical undergraduates, which would be more heterogeneous in these respects; for instance, due to their involvement in any (typically casual) employment, as well as to any income difference. Another

limitation is that we do not analyse changes in student performance. While we do observe improvements in the *e-assessment* cohort (both in terms of mean mark and pass rate), we choose not to include these data in this study as a key change was introduced in the summative assessment conditions that prevents a fair comparison. In particular, the *no e-assessment* exam was 'closed-book' and conducted traditionally with pen and paper, while the *e-assessment* exam was 'open-book/open-web' and on the Numbas system. Because of this change in exam conditions, the type and difficulty of the exam questions were adapted accordingly, thus introducing substantial differences; therefore, it would be arguably questionable to compare the two cohorts in this respect. Finally, an unavoidable overall limitation is that our study is characterized by many specific features, such as the nature of the module, its delivery, the teaching approach and quality, the e-assessment system adopted, etc.; therefore, as with similar studies, care must be taken before making any generalization.

CONCLUSION

We investigated the impact of introducing formative e-assessment on student attendance in a second year introductory statistics module at the University of the West of England (UWE Bristol). We used the Numbas e-assessment system, which offers key advantages over a traditional pen and paper approach, such as interaction (students engage online and receive automatic and instant feedback) and personalization (randomization makes assessments unique for each student). Upon introducing e-assessment, the average student attendance per session was found to increase by 10.2% in absolute terms and by 19.5% in relative terms. Statistical analysis provides strong evidence that such effect is significant. Importantly, the observed attendance increase is large enough to be of practical value. Student evaluation comments are consistent with an expectation that, compared to traditional pen and paper approaches, e-assessment can prove more authentic and engaging.

REFERENCES

Andrietti, V. (2014). Does lecture attendance affect academic performance? Panel data evidence for introductory macroeconomics. *International Review of Economics Education*, *15*, 1–16.

Angus, S.D., & Watson, J. (2009). Does regular online testing enhance student learning in the numerical sciences? Robust evidence from a large data set. *British Journal of Educational Technology*, *40*(2), 255–272.

Azab, E., Saksena, Y., Alghanem, T., Midle, J.B., Molgaard, K., Albright, S., & Karimbux, N. (2016). Relationship Among Dental Students' Class Lecture Attendance, Use of Online Resources, and Performance. *Journal of Dental Education*, *80*(4), 452–458.

Buchele, S. (2021). Evaluating the link between attendance and performance in higher education: The role of classroom engagement dimensions. *Assessment & Evaluation in Higher Education*, *46*(1), 132–150.

Butler-Henderson, K., & Crawford, J. (2020). A systematic review of online examinations: A pedagogical innovation for scalable authentication and integrity. *Computers & Education*, *159*, 104024.

Carroll, T., Casey, D., Crowley, J., Mulchrone, K., & Ní Shé, I. (2017). Numbas as an engagement tool for first-year Business Studies students. *MSOR Connections*, *15*(2), 42–50.

Casey, D., & Crowley, J. (2019). Embeddedness of mathematics e-assessment and attitudes affecting adoption. *MSOR Connections*, *17*(3).

Coates, H. (2005). The value of student engagement for higher education quality assurance. *Quality in Higher Education*, *11*(1), 25–36.

Credé, M., Roch, S.G., & Kieszczynka, U.M. (2010). Class Attendance in College: A meta-analytic review of the relationship of class attendance with grades and student characteristics. *Review of Educational Research*, *80*(2), 272–295.

Dey, I. (2018). Class attendance and academic performance: A subgroup analysis. *International Review of Economics Education*, *28*, 29–40.

Edwards, M.R., & Clinton, M.E. (2019). A study exploring the impact of lecture capture availability and lecture capture usage on student attendance and attainment. *Higher Education*, *77*(3), 403–421.

Foster, B., Perfect, C., & Youd, A. (2012). A completely client-side approach to e-assessment and e-learning of mathematics and statistics. *International Journal of e-assessment*, *2*(2).

Gillett-Swan, J. (2017). The challenges of online learning: Supporting and engaging the isolated learner. *Journal of Learning Design*, *10*(1), 20–30.

Graham, C. (2020). Assessment of computing in the mathematics curriculum using Numbas. *MSOR Connections*, *18*(2).

Holmes, N. (2015). Student perceptions of their learning and engagement in response to the use of a continuous e-assessment in an undergraduate module. *Assessment & Evaluation in Higher Education*, *40*(1), 1–14.

Holmes, N. (2018). Engaging with assessment: Increasing student engagement through continuous assessment. *Active Learning in Higher Education*, *19*(1), 23–34.

Kauffman, C.A., Derazin, M., Asmar, A., & Kibble, J.D. (2018). Relationship between classroom attendance and examination performance in a second-year medical pathophysiology class. *Advances in Physiology Education*, *42*(4), 593–598.

Kuh, G.D., Kinzie, J., Buckley, J.A., Bridges, B.K., & Hayek, J.C. (2011). Piecing together the student success puzzle: Research, propositions, and recommendations. *ASHE Higher Education Report* (Vol. 116). John Wiley & Sons.

Kuikka, M., Kitola, M., & Laakso, M-J. (2014). Challenges when introducing electronic exam. *Research in Learning Technology*, *22*.

Matthiasdottir, A., & Arnalds, H. (2016). E-Assessment: Students' Point of´ View. In *Proceedings of the 17th International Conference on Computer Systems and Technologies 2016* (pp. 369–374). CompSysTech '16, New York, NY, USA. Association for Computing Machinery.

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

Nordmann, E., Calder, C., Bishop, P., Irwin, A., & Comber, D. (2019). Turn up, tune in, don't drop out: The relationship between lecture attendance, use of lecture recordings, and achievement at different levels of study. *Higher Education*, *77*(6), 1065–1084.

Pagram, J., Cooper, M., Jin, H., & Campbell, A. (2018). Tales from the exam room: Trialing an e-exam system for computer education and design and technology students. *Education Sciences*, *8*(4), 188.

Perfect, C. (2015). A demonstration of Numbas, an e-assessment system for mathematical disciplines. In *CAA Conference*.

Prisacari, A.A, & Danielson, J. (2017). Rethinking testing mode: Should I offer my next chemistry test on paper or computer? *Computers & Education*, *106*, 1–12.

Rapanta, C., Botturi, L., Goodyear, P., Gu`ardia, 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*(3), 923–945.

Rolim, C., & Isaias, P. (2019). Examining the use of e-assessment in higher education: Teachers and students' viewpoints. *British Journal of Educational Technology*, *50*(4), 1785–1800.

Tetteh, G.A. (2018). Effects of classroom attendance and learning strategies on the learning outcome. *Journal of International Education in Business* .

Watermeyer, R., Crick, T., Knight, C., & Goodall, J. (2020). COVID-19 and digital disruption in UK universities: Afflictions and affordances of emergency online migration. *Higher Education*, pp. 1–19.

Weir, I., Gwynllyw, R., & Henderson, K. (2021). A case study in the e-assessment of statistics for non-specialists. *Journal of University Teaching & Learning Practice*, *18*(2), 05.

Williams, J.B., & Wong, A. (2009). The efficacy of final examinations: A comparative study of closed-book, invigilated exams and open-book, open-web exams. *British Journal of Educational Technology*, *40*(2), 227–236.