South American Publishing

SAP Gamification and Augmented Reality. 2023;2:12
doi: 10.62486/gr202312

ORIGINAL

Early assessment submission and course outcomes among students of an open online university

Entrega temprana de evaluaciones y resultados del curso en estudiantes de una universidad abierta en línea

Dhamelis Anaya Rios[1]
Analía Ortigoza[2]

[1] Universidade Federal do Rio Grande do Sul. Brazil
[2] Universidad Nacional de Tucumán, Tucumán, Argentina

Submitted: 15-04-2023 | Revised: 21-06-2023 | Accepted: 26-07-2023 | Published: 27-07-2023

Corresponding Author: Dhamelis Anaya Rios

ABSTRACT

Introduction: Students who withdraw or fail online courses often disengage from graded work early. It is uncertain whether early submission of continuous assessments predicts completion and passing, and whether this depends on prior study experience. Methods: Retrospective cohort of registrations in the Open University Learning Analytics Dataset (presentations from 2013 and 2014). Students still registered at day 60 were classified by submission of at least one tutor-marked or computer-marked assessment by that day. Logistic regression with module-presentation fixed effects gave adjusted odds ratios and standardized risk differences, with interaction tests by previous attempts and education. Results: The cohort comprised 23,943 students, 91.8% of whom submitted early. Completion was 85.3% with and 72.6% without early submission (adjusted odds ratio 1.30, 95% CI 1.10-1.53); passing was 62.4% and 7.9% (adjusted odds ratio 4.06, 95% CI 3.36-4.89). The adjusted risk difference for completion was larger with previous attempts (6.9 vs 2.9 percentage points) and rose with education. At a 90-day landmark, the association with completion included no effect. Conclusions: Early submission was strongly associated with passing and modestly with completion, and both varied with prior study experience. Because assessments count towards the mark, early submission is better used as a signal for outreach than as a causal lever.

Keywords: learning analytics; student engagement; educational assessment; online learning; academic success.

RESUMEN

Introducción: Los estudiantes que abandonan o reprueban cursos en línea suelen desvincularse temprano de las tareas calificadas. No está claro si la entrega temprana de evaluaciones continuas predice la finalización y la aprobación, ni si esto depende de la experiencia previa de estudio. Métodos: Cohorte retrospectiva de matrículas del Open University Learning Analytics Dataset (presentaciones de 2013 y 2014). Los estudiantes matriculados al día 60 se clasificaron según su entrega de al menos una evaluación calificada. La regresión logística con efectos fijos de módulo y presentación estimó razones de momios ajustadas y diferencias de riesgo estandarizadas, con pruebas de interacción por intentos previos y educación. Resultados: La cohorte incluyó 23 943 estudiantes, 91,8% con entrega temprana. La finalización fue 85,3% con y 72,6% sin entrega temprana (razón de momios ajustada 1,30; IC 95%: 1,10-1,53); la aprobación fue 62,4% y 7,9% (razón de momios ajustada 4,06; IC 95%: 3,36-4,89). La diferencia de riesgo ajustada para la finalización fue mayor con intentos previos (6,9 frente a 2,9 puntos porcentuales) y aumentó con la educación. Al día 90, la asociación con la finalización incluyó la ausencia de efecto. Conclusiones: La entrega temprana se asoció con fuerza a la aprobación y de forma modesta a la finalización, y ambas variaron con la experiencia previa. Como las evaluaciones cuentan para la nota, la entrega temprana sirve más como señal para el acompañamiento que como palanca causal.

Palabras clave: analítica del aprendizaje; compromiso estudiantil; evaluación educativa; aprendizaje en línea; éxito académico.

© 2023; The Authors. This Open access article is distributed under CC BY 4.0 licence.

INTRODUCTION

Withdrawal and failure are the best documented problems of online provision. Khaldi et al.(1) reviewed gamified e-learning in higher education and noted that digital learning environments present significant dropout and low completion rates, which they attributed mainly to weak motivation and engagement. Coffrin et al.(2) examined two massive open online courses with different curriculum and assessment designs and found that both lost most of their initially interested students, so that the courses looked alike on common indicators. Lee and Choi (3) had already modelled predictors of retention in online learning with a structural equation model, which placed the question of who stays at the centre of the field. The Open University Learning Analytics Dataset gave that question an open empirical base. Kuzilek et al.(4) released records of 32,593 students in 22 course presentations, with demographic data, assessment results and daily summaries of clicks in the virtual learning environment. Alhakbani and Alnassar (5) later reviewed the benchmark studies built on it. Most of those studies optimised prediction from activity logs: Aljohani et al.(6) trained a long short-term memory network on clickstreams, He et al.(7) combined static and sequential behaviour in a recurrent network, Liu et al.(8) compared weekly and monthly click features, and Jawad et al.(9) balanced classes with a resampling technique before fitting a random forest.

A second group of authors turned from clicks to graded work. Esteban et al.(10) examined whether assignment information improves the prediction of performance in distance learning, and Alshabandar et al.(11) asked whether performance and engagement in earlier computer-marked assessments affect achievement in later ones. Gupta et al.(12) built interpretable hidden Markov models on light-weight assessment features for early warning, and Jawthari and Stoffa (13) combined weekly engagement with assessment scores. Hlosta et al.(14) proposed a method that uses information on already submitted assessments to identify students at risk of failing a new course without legacy data, and Herodotou et al.(15) reported that teachers who used a predictive system aimed at students at risk of not submitting their next assignment had students with better outcomes. Studies outside that dataset reached similar designs with different data. Waheed et al.(16) built a neural network for the early identification of learners at risk in self-paced education, and the same group (17) predicted academic performance from assessment submission in a virtual learning environment. Olive et al.(18) trained networks to predict, two days before each due date, which students would submit their assignments on time, using 78,722 enrolments from five datasets, and Wan et al.(19) transferred a model between iterations of a hybrid course to design interventions. Al-azazi and Ghurab (20) proposed a day-wise multi-class model for massive open online courses so that instructors could intervene before the end of the course, and Umer et al.(21) added process-mining features to weekly interaction data.

The educational literature explains why submission behaviour matters. Black and Wiliam (22) reviewed the evidence that formative assessment raises achievement, and Hattie and Timperley (23) described feedback as one of the strongest influences on learning. Bulut et al.(24) argued that data from online formative assessments, rather than complex event logs, should be the foundation of predictive learning analytics, and King (25) found that usage of online formative tests was associated with examination performance after adjustment. Work on game elements, which Deterding et al.(26) defined as the use of game design elements in non-game contexts, points in the same direction: Huber et al.(27) found that game elements reduced attrition in an online learning task, mainly among participants with low self-efficacy, and Huang et al.(28) showed in a randomized field experiment that informing students about the completion rate of assignments increased the probability that they completed them.

Several questions remain open after these studies. The authors cited above reported accuracy, importance rankings or intervention effects, but none of them expressed the contribution of a single early behaviour, submitting a first assessment, as an absolute difference in completion and passing that a course team could act on. Jovanović et al.(29) showed that student conditions matter in predicting academic success, and Dirzyte et al.(30) observed that learners who finished an online programming course had higher initial knowledge scores than those who left after the first assessment, yet neither stratified the early-submission signal by prior study experience. Saqr et al.(31) added that predictors of success do not always port across learning designs, which makes a pooled estimate with design-specific adjustment preferable to a single-course result. The objective of this study was to estimate the association between early submission of continuous assessments and course completion and passing in the Open University Learning Analytics Dataset, and to test whether the association differs by prior study experience, defined by previous attempts at the module and by highest prior education.

METHODS

The study followed a retrospective cohort design with a landmark at day 60 of each course presentation and was reported according to the STROBE statement, with the transparency items of RECORD for routinely collected data. The source was the Open University Learning Analytics Dataset described by Kuzilek et al.,(4)an anonymised release under a CC BY 4.0 licence that covers seven modules delivered in 22 presentations that started in 2013 and 2014 (February and October starts, labelled B and J). The files were obtained programmatically from the distribution of the dataset in the UCI Machine Learning Repository (https://archive.ics.uci.edu/dataset/349/open+university+learning+analytics+dataset) and comprise the tables of student information, registration, assessments, student assessment results and daily platform clicks. The dataset contained 32,593 registrations of 28,785 distinct students; a student who took several modules or repeated one appears once per registration.

The assessment table lists, for each presentation, the type (tutor-marked assessment, computer-marked assessment or final examination), the due date in days from the start and the weight towards the final mark; the student assessment table records the day of each submission, the score and a flag for banked results. Platform activity was available as the number of clicks per student, site and day. Before the analysis, the tables were checked for duplicated registrations (0 were found), for registrations without a registration date (45 were found) and for the coding of the final result, and 1,909 banked submissions were identified.

The unit of analysis was the registration in a module presentation. Following the approach of authors who predicted outcomes from information available during a course,(9,14) a landmark was fixed at day 60 after the official start, roughly a quarter of the length of a presentation. Registrations entered the cohort if the presentation had at least one non-exam assessment due on or before day 60, if the student had registered on or before day 60, and if the student had not withdrawn on or before day 60. Three presentations of the one module whose first assessment fell after day 60 were therefore excluded, and the exclusions are counted in the Results. The restriction to students still enrolled at the landmark prevents classifying a student as non-submitting only because the student had already left.

The exposure was early assessment submission, defined as at least one tutor-marked or computer-marked assessment submitted on or before day 60. Final examinations were not counted. Results carried over from a previous attempt (banked assessments) were not counted as submissions in the current presentation. These assessments are graded and contribute to the final mark, but individually their weight is small and they are designed to give students feedback during the course, which is why they are treated here as continuous assessments with a formative function rather than as summative tests. Exposure was also coded as none, one, or two or more early submissions.

Two outcomes were taken from the final result recorded by the university. Completion was any final result other than withdrawn, and passing was a final result of pass or distinction, so that withdrawn and failed students were non-passers. Prior study experience was represented by two effect modifiers: previous attempts at the same module (none versus one or more) and highest education before enrolment, grouped as below A level (including no formal qualifications), A level or equivalent, and higher education (a university or postgraduate qualification). Covariates were module presentation, gender, age band, deprivation band of the area of residence (with an unknown category), declared disability, studied credits per 30 credits, highest education, previous attempts, the natural logarithm of one plus the platform clicks accumulated up to day 60, and registration day relative to the start in units of 30 days.

The protocol for the main models was written before they were fitted and is described here in full; the sensitivity analysis on banked results was added after the first results showed unexpected pass rates among repeating students. Proportions were compared with the chi-square test. For each outcome, a logistic regression with module-presentation fixed effects and heteroskedasticity-robust standard errors gave the adjusted odds ratio for early submission with its 95% confidence interval. Because odds ratios are hard to interpret when the outcome is common, the adjusted risk difference was obtained by standardization: predicted probabilities were computed for every student with the exposure set to one and to zero, and the mean difference was reported in percentage points with a percentile 95% confidence interval from 300 bootstrap resamples (200 in the sensitivity analysis restricted to students without banked results). Effect modification was tested by comparing nested models with and without the product term by a likelihood-ratio test, and the four interaction p values (two modifiers, two outcomes) were adjusted with the Holm procedure. Stratum-specific estimates came from models fitted within each stratum.

Sensitivity analyses repeated the main model with the exposure restricted to tutor-marked assessments, with students of unknown deprivation band removed, with a landmark at day 90 that included all 22 presentations, and with students who had any banked result removed. The incremental discrimination of the early-submission count was assessed with five-fold stratified cross-validation (seed 42) of logistic models for three nested predictor sets: demographic and module variables, the same plus log clicks, and the same plus the number of early submissions. The area under the receiver operating characteristic curve was reported with a 95% bootstrap interval of 200 resamples, and the paired difference between the last two models was bootstrapped in the same way. All tests were two-sided with a significance level of 0.05; with this number of registrations, differences of a few percentage points were detectable, so effect sizes are given priority over p values.

The dataset contains no direct identifiers and was released by the university for open research use, so no new data were collected from participants and no additional ethical approval was sought for this secondary analysis. No generative model was used to code or classify any record; every figure comes from the analysis script.

RESULTS

Of the 32,593 registrations in the dataset, 23,943 formed the analytic cohort. Presentations without an assessment due by day 60 removed 2,534 registrations, withdrawal on or before day 60 removed 6,101, and registration after day 60 or without a registration date removed 15. The cohort covered 19 presentations of 6 modules.

By day 60, 21,982 students (91.8%) had submitted at least one assessment and 1,961 had not; 4,608 of the early submitters had sent one assessment and 17,374 had sent two or more. Early submission was less frequent among students with previous attempts (77.6%) than among those without (94.0%), and the median number of clicks before day 60 was 455 in the early group and 31 in the other (Table 1).

Table 1. Characteristics and outcomes of eligible students by early assessment submission.

Characteristic	Without early submission	With early submission	All eligible
Students, n	1,961	21,982	23,943
Female, n (%)	751 (38.3)	9,191 (41.8)	9,942 (41.5)
Age 35 years or older, n (%)	560 (28.6)	6,532 (29.7)	7,092 (29.6)
One or more previous attempts, n (%)	724 (36.9)	2,502 (11.4)	3,226 (13.5)
Below A level education, n (%)	989 (50.4)	7,985 (36.3)	8,974 (37.5)
A level education, n (%)	702 (35.8)	10,070 (45.8)	10,772 (45.0)
Higher education, n (%)	270 (13.8)	3,927 (17.9)	4,197 (17.5)
Declared disability, n (%)	245 (12.5)	1,877 (8.5)	2,122 (8.9)
Most deprived 20% of areas, n (%)	493 (25.1)	4,121 (18.7)	4,614 (19.3)
Studied credits, median (IQR)	60 (60-120)	60 (60-90)	60 (60-100)
Platform clicks before day 60, median (IQR)	31 (3-101)	455 (234-866)	411 (193-813)
Completed the course, n (%)	1,423 (72.6)	18,751 (85.3)	20,174 (84.3)
Passed the course, n (%)	155 (7.9)	13,712 (62.4)	13,867 (57.9)

Note: Early submission means at least one tutor-marked or computer-marked assessment submitted on or before day 60. IQR = interquartile range. The most deprived 20% corresponds to the two lowest bands of the index of multiple deprivation.

In the whole cohort, 20,174 students (84.3%) completed the course and 13,867 (57.9%) passed, including 2,628 with distinction. Completion reached 85.3% among early submitters and 72.6% among the rest (chi-square = 220.2, p < 0.001). Passing reached 62.4% and 7.9%, respectively (chi-square = 2,192.0, p < 0.001) (Table 2).

After adjustment, the odds of completion for early submitters were 1.30 times those of non-submitters (95% CI 1.10-1.53), equivalent to 3.5 percentage points (95% CI 0.9-6.3). The odds of passing were 4.06 times as high (95% CI 3.36-4.89), equivalent to 27.9 percentage points (95% CI 24.6-31.0). With only module presentation in the model, the odds ratios were 2.21 for completion and 19.9 for passing (Table 2).

The gradient with the number of early submissions was uneven for completion. One early submission gave an adjusted odds ratio of 0.92 (95% CI 0.78-1.10) and two or more gave 1.76 (95% CI 1.49-2.09). For passing, the corresponding ratios were 1.94 (95% CI 1.58-2.38) and 6.98 (95% CI 5.67-8.59).

Table 2. Completion and passing by early assessment submission, overall and by prior study experience.

Outcome	Stratum	Students, n	Without early submission, %	With early submission, %	Adjusted OR (95% CI)	Adjusted risk difference, percentage points (95% CI)	P for interaction
Completion	All eligible students	23,943	72.6	85.3	1.30 (1.10 to 1.53)	3.5 (0.9 to 6.3)	
Completion	No previous attempt	20,717	70.5	85.7	1.24 (1.02 to 1.52)	2.9 (0.3 to 5.4)	<0.001
Completion	One or more previous attempts	3,226	76.1	82.5	1.55 (1.16 to 2.08)	6.9 (1.8 to 11.4)	
Completion	Below A level	8,974	74.4	82.6	1.11 (0.87 to 1.42)	1.5 (-1.8 to 5.5)	<0.001
Completion	A level	10,772	71.8	86.8	1.38 (1.06 to 1.81)	4.1 (0.3 to 8.3)	
Completion	Higher education	4,197	67.8	87.1	1.94 (1.27 to 2.95)	9.0 (2.7 to 16.5)	
Pass	All eligible students	23,943	7.9	62.4	4.06 (3.36 to 4.89)	27.9 (24.6 to 31.0)	
Pass	No previous attempt	20,717	2.0	64.2	19.40 (12.82 to 29.37)	50.5 (46.3 to 54.1)	<0.001
Pass	One or more previous attempts	3,226	18.0	47.9	1.43 (1.11 to 1.84)	7.1 (2.0 to 11.5)	
Pass	Below A level	8,974	4.7	53.6	4.60 (3.32 to 6.37)	28.0 (23.3 to 33.0)	0.428
Pass	A level	10,772	10.3	66.2	4.01 (3.03 to 5.31)	28.8 (23.6 to 34.5)	
Pass	Higher education	4,197	13.7	70.5	3.28 (2.16 to 4.97)	24.2 (16.5 to 33.0)	

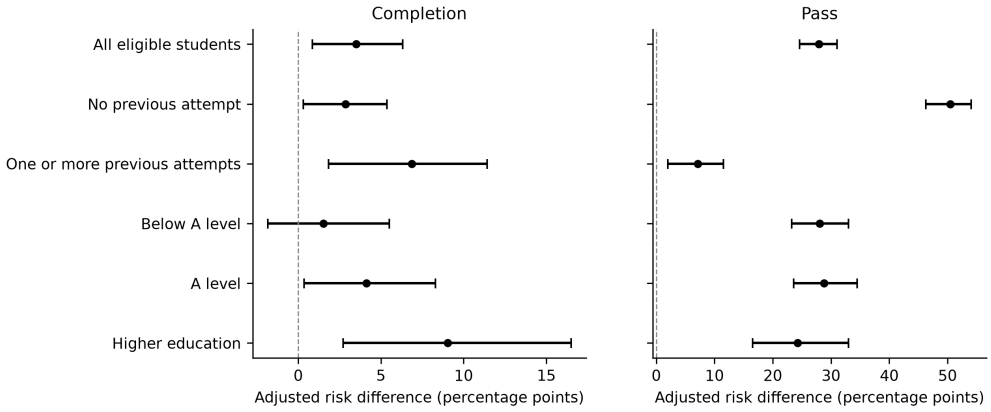
Note: OR = odds ratio; CI = confidence interval. Estimates are adjusted for module presentation, gender, age band, deprivation band, disability, studied credits, highest education, previous attempts, platform clicks up to day 60 and registration day, except that the stratifying variable is omitted within its own strata. The risk difference is the standardized mean difference in predicted probability. The p value of the interaction applies to the group of rows that follows it.

Prior study experience modified the association with completion. The adjusted risk difference was 6.9 percentage points (95% CI 1.8-11.4) among students with previous attempts and 2.9 (95% CI 0.3-5.4) among those without (interaction $p < 0.001$). It increased with education, from 1.5 percentage points (95% CI −1.8 to 5.5) below A level to 9.0 (95% CI 2.7-16.5) in higher education (interaction $p < 0.001$) (Table 2, Figure 1).

The pattern reversed for passing in relation to previous attempts. The adjusted risk difference was 50.5 percentage points (95% CI 46.3-54.1) in students without previous attempts and 7.1 (95% CI 2.0-11.5) in those with them (interaction $p < 0.001$). Among non-submitters, 2.0% of students without previous attempts passed against 18.0% of those with them. The risk difference did not vary with education, at 28.0, 28.8 and 24.2 percentage points (interaction $p = 0.428$). All three interactions that were below 0.05 remained so after the Holm adjustment.

Results carried over from earlier attempts affected the passing comparison. Banked assessments existed for 480 students of the cohort, all of them with previous attempts, and for 418 of the 724 repeating students without an early submission (57.7%). In that last group, 29.9% of those with banked results passed and 1.6% of those without them did. After removing the students with banked results (23,463 remained), the adjusted risk difference for passing was 36.5 percentage points in repeating students and 50.5 in the others (interaction $p = 0.216$), whereas the interaction for completion remained ($p = 0.011$), with 8.3 and 2.9 percentage points.

Figure 1. Adjusted risk difference for early assessment submission by prior study experience.



Note: Points are standardized adjusted risk differences in percentage points and bars are 95% bootstrap confidence intervals. The dashed line marks no difference.

Early submission improved discrimination only slightly once platform activity was known. For completion, the cross-validated area under the curve was 0.629 for demographic and module variables, 0.650 after adding clicks and 0.668 after adding the number of early submissions, a gain of 0.017 (95% CI 0.013-0.021). For passing, the values were 0.659, 0.764 and 0.784, a gain of 0.020 (95% CI 0.017-0.023).

The main estimates changed little when the exposure was limited to tutor-marked assessments (adjusted odds ratio 1.18 for completion, 95% CI 1.03-1.35, and 2.38 for passing, 95% CI 2.11-2.68) or when the 990 students of unknown deprivation band were removed. The landmark at day 90 included 25,551 students from all 22 presentations, and 91.9% had submitted by then. The adjusted odds ratio for completion was 1.07 (95% CI 0.89-1.30), which includes no association, whereas that for passing was 4.95 (95% CI 4.08-5.99). The interaction with previous attempts remained below 0.001 for both outcomes at day 90.

DISCUSSION

The main finding is a contrast between two outcomes that are often treated as one. Students who had submitted a graded assessment by day 60 passed far more often, whereas their advantage in completion was modest, depended on the choice of landmark and was concentrated among students with more education and with previous attempts. Pooled across six modules, this is in line with Esteban et al.,(10)who reported that assignment information improves the prediction of performance in distance learning, and with Alshabandar et al.,(11)who found that performance and engagement in earlier assessments relate to later achievement. The pass difference found here is large, but it should not be read as the effect of the behaviour itself, because a student who submits nothing for two months has little route to a passing mark.

That mechanical link explains why passing and completion diverge. Just as Bulut et al.(24) argued, data from graded online tasks are close to the outcome that learning analytics tries to predict, so their predictive value is partly definitional. King(25)observed an association between formative test usage and examination performance in an accounting module, and the present estimates extend that association to a whole open university, but neither study can separate the learning benefit of practice from the selection of diligent students. Completion is less affected by this circularity, since a student can stay registered without submitting, and the smaller association we found for it is the more informative estimate about engagement. Unlike the findings of Esteban et al.,(10)the contribution of the early-submission count to discrimination was small once clicks were in the model, a gain of about 0.02 in the area under the curve, which suggests that early activity in the virtual learning environment already carries most of the prognostic information and that assessment submission refines rather than replaces it, as Jawthari and Stoffa(13)also combined both sources.

The sensitivity of completion to the landmark deserves attention. At day 60 the odds ratio for completion was clearly above one, but at day 90, when more students had submitted something and the cohort retained students who left between day 60 and day 90, it was not distinguishable from one. Coffrin et al.(2)noted that attrition in open courses is high and similar across designs, and

Nguyen et al.,(32)using a mixed-effect logistic regression on records of the Open University itself, examined how study breaks and preparation weeks affected the odds of passing a module, so the weeks in which a landmark is placed may carry information of their own. A plausible reading is that the first weeks separate students who never engage from those who do, while later withdrawal reflects events that early submission does not capture, such as workload or life circumstances. Lee and Choi(3)modelled retention with a structural equation model, and whether their predictors explain withdrawal after day 60 was not tested here.

Prior study experience changed the picture in ways that are partly expected and partly an artefact of the dataset. In keeping with Pi et al.,(33)who reported that prior knowledge moderated the effects of motivation compositions on interaction in video-based learning, the absolute completion advantage grew with education, from 1.5 percentage points below A level to 9.0 in higher education. Dirzyte et al.(30)reported that learners with higher initial knowledge scores finished an online programming course more often than those who left after the first assessment, and Van Campenhout et al.(34)accounted for prior knowledge when they tested the doer effect across seven courses; our result suggests that students with more schooling convert an early submission into persistence more readily, although the broad interval in the lowest education group allows a smaller difference. The reversal for passing among repeating students has a simpler cause. Their results from earlier attempts can be banked and counted towards the final mark, so some of them pass without early submissions in the current presentation; when students with banked results were removed, the difference in passing between repeating and other students was no longer distinguishable, whereas the difference in completion persisted.

The literature on portability argues for caution before transferring any of these figures. Saqr et al.(31)and Jovanović et al.(29)found that predictors of success vary with course design and with student conditions, and Wang and Mousavi,(35)in a meta-analysis of 88 studies, reported that not every log variable is related to achievement and that relationships varied across courses. Quick et al.(36)replicated a temporal modelling study over 19 courses and found that only some general patterns of self-regulated learning generalised. The use of module-presentation fixed effects here guards against baseline differences between courses, but the stratified estimates still pool six modules with different assessment calendars, and one module with no assessment before day 60 had to be excluded from the main analysis. Transfer to Latin America should likewise be argued rather than assumed. Aguirre Montano and Cabrera-Loayza(37)predicted dropout early in online courses in Ecuador from neural networks, Queiroga et al.(38)built risk models from platform interactions in a Brazilian distance technical course, and Battaglin et al.(39)compared algorithms on a small Brazilian programming course; all three worked with data of regional institutions that differ from the Open University in size and assessment structure. The finding that a first submission is more informative for repeating students and for those with more education may therefore hold in regional universities with similar structures, but it needs local data.

The comparison with studies that used the same data on prediction accuracy is limited by the different purposes. Alshabandar et al.,(11)Jawad et al.(9)and Liu et al.(8)evaluated classifiers on held-out students, and our cross-validated areas under the curve, which were lower for completion than for passing, are not directly comparable with their metrics because the cohort is restricted to students still enrolled at day 60 and because a landmark at day 60 leaves less variation in the outcome than a prediction at registration. Hlosta et al.(14)targeted a new course without legacy data and used the assessments already submitted, which is the situation of a course team in its first week; our pooled estimates suggest that such a method would find most of its signal in whether any assessment is submitted, with a smaller additional contribution from how many.

Fairness is a practical limit on using such signals. Al-Zawqari and Vandersteen (40) showed that hiding protected attributes lowered the fairness of models for online STEM courses, and Deho et al.(41) reported that demographic imbalance in training data can produce unfair models. Jawthari and Stoffa (42) examined whether at-risk models should include protected features. Here gender, age band, deprivation and disability were covariates, not predictors for triage, and the early-submission group differed from the rest in disability and deprivation, so a rule that contacts only non-submitters would reach a group that is also more disadvantaged, which is a reason to combine it with support rather than sanction.

The size of the change when covariates were added also needs explanation. With module presentation alone, the odds ratio for passing was close to 20, and it fell to about 4 after adjustment for demographic variables, previous attempts and platform clicks, while the odds ratio for completion fell from 2.21 to 1.30. Aljohani et al.(6)and Liu et al.(8)built their models on clickstreams, and our results show why those models perform well: students who submit early are the same students who click, and most of the crude association is carried by activity. Qiu et al.(43)argued

that e-learning behaviours are inherently correlated and proposed to classify and fuse them before prediction, which is consistent with the shared variance seen here. He et al.(7)combined static biographical data with sequential behaviour for the same reason. What remains after adjustment is the part of early submission that clicks do not explain, and its size, a gain of about two points of the area under the curve and an odds ratio of 4.06 for passing, is the figure that matters for a team deciding whether to monitor submissions in addition to logins.

The dose pattern qualifies the binary exposure. For completion, a single early submission was not associated with a higher probability (adjusted odds ratio 0.92), whereas two or more were. A reading in line with Ruipérez-Valiente et al.,(44)who built early predictions of certificate attainment with variable importance in a massive open online course, is that a lone submission may reflect a student who tried once and stopped, and that sustained submission is the marker of persistence. Gupta et al.(12)reached a related conclusion in the same dataset by modelling the sequence of light-weight assessments, not a single event. For passing, even one submission was associated with roughly double the odds, which fits the fact that any graded score moves a student away from a zero total. In practice, the threshold for outreach could therefore differ by outcome: no submission by the first deadline for passing, and fewer than two by day 60 for completion.

What could course teams do with this signal? Herodotou et al.(15)found that teachers who intervened with students flagged as unlikely to submit their next assignment, using the Open University predictive system, had students who performed better, and Huang et al.(28)showed that a descriptive norm about assignment completion increased the probability of completing, which is exactly the behaviour at stake here. Gamified designs offer a second route. Huber et al.(27)found that game elements reduced attrition, particularly among learners with low self-efficacy, and Leung et al.(45)reported that gamification in a massive open online course helped when it gave personalised feedback matched to the learner's goal orientation. Puig et al.(46)designed a dynamic adaptive gamification approach to counter low retention, and Major and Mira da Silva(47)concluded from a review that the balance between game elements and learning is the main design problem. Our data cannot say whether such elements would change behaviour, but they identify who to target and when: the first weeks, and especially students who have studied the module before, among whom early submission was least common (77.6% against 94.0%). Henderson et al.(48) observed that students' experience of feedback usefulness differs between online and international students and by year of study, and a first assessment that returns feedback quickly would give such students a reason to continue. Nieberding and Heckler (49) distinguished classes of students by the time before the deadline at which they submitted weekly work and linked later submission to lower grades, which suggests recording submission timing in future work.

Several strengths and limitations should be weighed. The data are open, the cohort is large and the landmark design avoids classifying leavers as non-submitters. The assessments, however, are part of the grade, so the pass association is partly built into the outcome, and a design that separates formative activity from graded activity would be needed to isolate learning; ungraded practice or self-test items, which Van Campenhout et al.(34)studied through the doer effect, should be analysed in a dataset that records them. The observational design leaves confounding by motivation and available time, which no covariate here measures, so adjusted estimates remain associations and a randomized prompt to submit would be the way to test causality. Students who left before day 60 were excluded, nearly a fifth of registrations in the dataset, so the results apply to those still enrolled at the landmark, and a study of earlier leavers would need a landmark before the first deadline. Prior study experience was coded from two variables of the dataset that miss informal learning, which could be addressed by linking registrations across presentations. Finally, the data cover two years of a single distance-teaching university, and replication in other open datasets would indicate how far the stratified differences travel.

CONCLUSIONS

Early submission of continuous assessments, observed by day 60, was strongly associated with passing and modestly associated with completion among students still enrolled at that point; the association with completion was not present at a day-90 landmark. Prior study experience modified both associations: the absolute advantage in completion was larger for students with previous attempts and for those with higher education, whereas the absolute advantage in passing was much smaller among repeating students, a difference that disappeared once students with results carried over from earlier attempts were set aside, and did not vary with education. Early submission added only about two points of discrimination to platform activity, so its main use is

as a prompt for supportive outreach and for the design of first-assessment experiences, not as a causal lever or a stand-alone screening rule.

REFERENCES

1. Khaldi A, Bouzidi R, Nader F. Gamification of e-learning in higher education: a systematic literature review. Smart Learn Environ. 2023;10(1):10. doi: 10.1186/s40561-023-00227-z
2. Coffrin C, Corrin L, de Barba P, Kennedy G. Visualizing patterns of student engagement and performance in MOOCs. Proceedings of the Fourth International Conference on Learning Analytics And Knowledge. 2014:83-92. doi: 10.1145/2567574.2567586
3. Lee Y, Choi J. A structural equation model of predictors of online learning retention. The Internet and Higher Education. 2013;16(1):36-42. doi: 10.1016/j.iheduc.2012.01.005
4. Kuzilek J, Hlosta M, Zdrahal Z. Open University Learning Analytics dataset. Sci Data. 2017;4(1):170171. doi: 10.1038/sdata.2017.171
5. Alhakbani HA, Alnassar FM. Open Learning Analytics: A Systematic Review of Benchmark Studies using Open University Learning Analytics Dataset (OULAD). 2022 7th International Conference on Machine Learning Technologies (ICMLT). 2022:81-86. doi: 10.1145/3529399.3529413
6. Aljohani NR, Fayoumi A, Hassan SU. Predicting At-Risk Students Using Clickstream Data in the Virtual Learning Environment. Sustainability. 2019;11(24):7238. doi: 10.3390/su11247238
7. He Y, Chen R, Li X, Hao C, Liu S, Zhang G, et al. Online At-Risk Student Identification using RNN-GRU Joint Neural Networks. Information. 2020;11(10):474. doi: 10.3390/info11100474
8. Liu Y, Fan S, Xu S, Sajjanhar A, Yeom S, Wei Y. Predicting Student Performance Using Clickstream Data and Machine Learning. Education Sciences. 2022;13(1):17. doi: 10.3390/educsci13010017
9. Jawad K, Shah MA, Tahir M. Students' Academic Performance and Engagement Prediction in a Virtual Learning Environment Using Random Forest with Data Balancing. Sustainability. 2022;14(22):14795. doi: 10.3390/su142214795
10. Esteban A, Romero C, Zafra A. Assignments as Influential Factor to Improve the Prediction of Student Performance in Online Courses. Applied Sciences. 2021;11(21):10145. doi: 10.3390/app112110145
11. Alshabandar R, Hussain A, Keight R, Khan W. Students Performance Prediction in Online Courses Using Machine Learning Algorithms. 2020 International Joint Conference on Neural Networks (IJCNN). 2020:1-7. doi: 10.1109/ijcnn48605.2020.9207196
12. Gupta A, Garg D, Kumar P. Mining Sequential Learning Trajectories With Hidden Markov Models For Early Prediction of At-Risk Students in E-Learning Environments. IEEE Trans Learning Technol. 2022;15(6):783-797. doi: 10.1109/tlt.2022.3197486
13. Jawthari M, Stoffa V. Predicting At-Risk Students Using Weekly Activities and Assessments. Int J Emerg Technol Learn. 2022;17(19):59-73. doi: 10.3991/ijet.v17i19.31349
14. Hlosta M, Zdrahal Z, Zendulka J. Ouroboros. Proceedings of the Seventh International Learning Analytics & Knowledge Conference. 2017:6-15. doi: 10.1145/3027385.3027449
15. Herodotou C, Hlosta M, Boroowa A, Rienties B, Zdrahal Z, Mangafa C. Empowering online teachers through predictive learning analytics. Brit J Educational Tech. 2019;50(6):3064-3079. doi: 10.1111/bjet.12853
16. Waheed H, Hassan SU, Nawaz R, Aljohani NR, Chen G, Gasevic D. Early prediction of learners at risk in self-paced education: A neural network approach. Expert Systems with Applications. 2023;213:118868. doi: 10.1016/j.eswa.2022.118868
17. Waheed H, Nisar I, Khalid MUN, Shahid A, Aljohani NR, Hassan SU, et al. Predicting Academic Performance of Students from the Assessment Submission in Virtual Learning Environment. In: Springer Proceedings in Complexity. Springer International Publishing; 2023. p. 417-424. doi: 10.1007/978-3-031-19560-0_33
18. Olive DM, Huynh DQ, Reynolds M, Dougiamas M, Wiese D. A Quest for a One-Size-Fits-All Neural Network: Early Prediction of Students at Risk in Online Courses. IEEE Trans Learning Technol. 2019;12(2):171-183. doi: 10.1109/tlt.2019.2911068
19. Wan H, Liu K, Yu Q, Gao X. Pedagogical Intervention Practices: Improving Learning Engagement Based on Early Prediction. IEEE Trans Learning Technol. 2019;12(2):278-289. doi: 10.1109/tlt.2019.2911284
20. Al-azazi FA, Ghurab M. ANN-LSTM: A deep learning model for early student performance prediction in MOOC. Heliyon. 2023;9(4):e15382. doi: 10.1016/j.heliyon.2023.e15382
21. Umer R, Susnjak T, Mathrani A, Suriadi S. On predicting academic performance with process mining in learning analytics. JRIT&L. 2017;10(2):160-176. doi: 10.1108/jrit-09-2017-0022
22. Black P, Wiliam D. Assessment and Classroom Learning. Assessment in Education: Principles, Policy & Practice. 1998;5(1):7-74. doi: 10.1080/0969595980050102
23. Hattie J, Timperley H. The Power of Feedback. Review of Educational Research. 2007;77(1):81-112. doi: 10.3102/003465430298487
24. Bulut O, Gorgun G, Yildirim-Erbasli SN, Wongvorachan T, Daniels LM, Gao Y, et al. Standing on the shoulders of giants: Online formative assessments as the foundation for predictive learning analytics models. Brit J Educational Tech. 2023;54(1):19-39. doi: 10.1111/bjet.13276
25. King D. Assessing the benefits of online formative assessments on student performance. JLDHE. 2023;2023(27). doi: 10.47408/jldhe.vi27.883
26. Deterding S, Dixon D, Khaled R, Nacke L. From game design elements to gamefulness. Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments. 2011:9-15. doi: 10.1145/2181037.2181040
27. Huber SE, Cortez R, Kiili K, Lindstedt A, Ninaus M. Game elements enhance engagement and mitigate attrition in online learning tasks. Computers in Human Behavior. 2023;149:107948. doi: 10.1016/j.chb.2023.107948

28. Huang N, Zhang J, Burtch G, Li X, Chen P. Combating Procrastination on Massive Online Open Courses via Optimal Calls to Action. Information Systems Research. 2021;32(2):301-317. doi: 10.1287/isre.2020.0974
29. Jovanović J, Saqr M, Joksimović S, Gašević D. Students matter the most in learning analytics: The effects of internal and instructional conditions in predicting academic success. Computers & Education. 2021;172:104251. doi: 10.1016/j.compedu.2021.104251
30. Dirzyte A, Perminas A, Kaminskis L, Žebrauskas G, Sederevičiūtė – Pačiauskienė Ž, Šliogerienė J, et al. Factors contributing to dropping out of adults' programming e-learning. Heliyon. 2023;9(12):e22113. doi: 10.1016/j.heliyon.2023.e22113
31. Saqr M, Jovanovic J, Viberg O, Gašević D. Is there order in the mess? A single paper meta-analysis approach to identification of predictors of success in learning analytics. Studies in Higher Education. 2022;47(12):2370-2391. doi: 10.1080/03075079.2022.2061450
32. Nguyen Q, Thorne S, Rienties B. How do students engage with computer-based assessments: impact of study breaks on intertemporal engagement and pass rates. Behaviormetrika. 2018;45(2):597-614. doi: 10.1007/s41237-018-0060-1
33. Pi Z, Zhang X, Zhang X, Gao M, Li X. Students' prior knowledge moderates the effects of group motivation compositions on learning from video lectures. Brit J Educational Tech. 2023;54(6):1814-1836. doi: 10.1111/bjet.13316
34. Van Campenhout R, Jerome B, Dittel JS, Johnson BG. The Doer Effect at Scale: Investigating Correlation and Causation Across Seven Courses. LAK23: 13th International Learning Analytics and Knowledge Conference. 2023:357-365. doi: 10.1145/3576050.3576103
35. Wang Q, Mousavi A. Which log variables significantly predict academic achievement? A systematic review and meta-analysis. Brit J Educational Tech. 2023;54(1):142-191. doi: 10.1111/bjet.13282
36. Quick JD, Motz B, Morrone A. Lost in Translation: Determining the Generalizability of Temporal Models across Course Contexts. LAK23: 13th International Learning Analytics and Knowledge Conference. 2023:273-283. doi: 10.1145/3576050.3576092
37. Aguirre Montano HS, Cabrera-Loayza MC. Early prediction of dropout in online courses using Artificial Neural Networks. 2020 XV Conferencia Latinoamericana de Tecnologias de Aprendizaje (LACLO). 2020:1-6. doi: 10.1109/laclo50806.2020.9381190
38. Queiroga EM, Lopes JL, Kappel K, Aguiar M, Araújo RM, Munoz R, et al. A Learning Analytics Approach to Identify Students at Risk of Dropout: A Case Study with a Technical Distance Education Course. Applied Sciences. 2020;10(11):3998. doi: 10.3390/app10113998
39. Battaglin R, Munoz R, Ramos VC, Cechinel C. Predicting at-risk students with LMS data: a comparison between Adaboost and LSTM algorithms. 2022 XVII Latin American Conference on Learning Technologies (LACLO). 2022:1-4. doi: 10.1109/laclo56648.2022.10013469
40. Al-Zawqari A, Vandersteen G. Fairness in Predictive Learning Analytics: A Case Study in Online STEM Education. 2023 IEEE Frontiers in Education Conference (FIE). 2023:1-5. doi: 10.1109/fie58773.2023.10343059
41. Deho OB, Joksimovic S, Liu L, Li J, Zhan C, Liu J. Assessing the Fairness of Course Success Prediction Models in the Face of (Un)equal Demographic Group Distribution. Proceedings of the Tenth ACM Conference on Learning @ Scale. 2023:48-58. doi: 10.1145/3573051.3593381
42. Jawthari M, Stoffa V. Should Course-based At-risk Predication Models Include Protected Features?. 2022 13th International Conference on Information and Communication Systems (ICICS). 2022:257-261. doi: 10.1109/icics55353.2022.9811127
43. Qiu F, Zhang G, Sheng X, Jiang L, Zhu L, Xiang Q, et al. Predicting students' performance in e-learning using learning process and behaviour data. Sci Rep. 2022;12(1):453. doi: 10.1038/s41598-021-03867-8
44. Ruipérez-Valiente JA, Cobos R, Muñoz-Merino PJ, Andujar Á, Delgado Kloos C. Early Prediction and Variable Importance of Certificate Accomplishment in a MOOC. In: Lecture Notes in Computer Science. Springer International Publishing; 2017. p. 263-272. doi: 10.1007/978-3-319-59044-8_31
45. Leung ACM, Santhanam R, Kwok RCW, Yue WT. Could Gamification Designs Enhance Online Learning Through Personalization? Lessons from a Field Experiment. Information Systems Research. 2023;34(1):27-49. doi: 10.1287/isre.2022.1123
46. Puig A, Rodríguez I, Rodríguez Á, Gallego I. Evaluating Learner Engagement with Gamification in Online Courses. Applied Sciences. 2023;13(3):1535. doi: 10.3390/app13031535
47. Major RR, Mira da Silva M. Gamification in MOOCs: A systematic literature review. Cogent Education. 2023;10(2):2275820. doi: 10.1080/2331186x.2023.2275820
48. Henderson M, Ryan T, Boud D, Dawson P, Phillips M, Molloy E, et al. The usefulness of feedback. Active Learning in Higher Education. 2021;22(3):229-243. doi: 10.1177/1469787419872393
49. Nieberding M, Heckler AF. Patterns in assignment submission times: Procrastination, gender, grades, and grade components. Phys Rev Phys Educ Res. 2021;17(1):013106. doi: 10.1103/physrevphyseducres.17.013106

FUNDING

No financing

CONFLICT OF INTEREST

None

AUTHORSHIP CONTRIBUTIONS

Drafting – original draft: Dhamelis Anaya Rios, Analía Ortigoza.
Writing–review and editing: Dhamelis Anaya Rios, Analía Ortigoza.