Studies in Higher Education

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/cshe20

How you teach and who you teach both matter: lessons from learning analytics data

Robert J. Summers, Adrian P. Burgess, Helen E. Higson & Elisabeth Moores

STUDIES IN HIGHER EDUCATION
https://doi.org/10.1080/03075079.2023.2245424

How you teach and who you teach both matter: lessons from learning analytics data

Robert J. Summers [a], Adrian P. Burgess [a], Helen E. Higson [b] and Elisabeth Moores [a]

[a]College of Health and Life Sciences, Aston University, Birmingham, UK; [b]Aston Business School, Aston University, Birmingham, UK

ABSTRACT
To explore potential effects of disadvantage on engagement and attainment under different teaching and assessment regimes, the influence of pedagogic changes implemented during the COVID-19 pandemic on attainment and engagement of students from different backgrounds were compared using a cohort-study design. Learner analytics and attainment data from first year undergraduate students during three learning regimes were compared: (i) in-person teaching and assessment, (ii) in-person teaching with online assessment, and (iii) online teaching and assessment. The gap in end-of-year mark between disadvantaged students and their peers was widest when teaching and assessment was online, with poorer outcomes for disadvantaged students, although the gap in the percentage of students passing all their modules did not change. Overall, online teaching and assessment during the pandemic was associated with a widening attainment gap between disadvantaged students and their peers. Possible explanations for this are discussed, including the relationship between attainment and engagement. Higher Education providers should monitor and review the potential implications of their chosen education strategy on different groups of students: how you teach and who you teach are both important.

ARTICLE HISTORY
Received 3 October 2022
Accepted 2 August 2023

KEYWORDS
Learning analytics; disadvantage; student engagement; higher education; digital poverty; attainment

Introduction

The COVID-19 pandemic forced many higher education providers to move much of their teaching and assessment online in 2020 and 2021, giving rise to serious concerns about the potential effects of digital poverty on education (i.e. exclusion from education due to lack of appropriate hardware, software, or internet connectivity). Using survey results, Rodríguez-Planas (2022) reported that retention rates fell – up to 34% of students considered dropping a course – and this rate was even higher in low-income students (see also Bird, Castleman, and Lohner 2022; Kofoed et al. 2021). In contrast, El Said (2021) reported no negative effects, and that the shift to online learning could be beneficial in terms of cost-effective expansion of provision. Similarly, Kizilcec, Makridis, and Sadowski (2021) reported higher engagement during the pandemic amongst existing online students from lower income regions. As the ongoing risks of COVID-19 appear to be reducing, higher education providers are considering what the 'new normal' should entail, commonly trying to harness some of the perceived advantages and flexibility of online learning, whilst still providing aspects of an

CONTACT Elisabeth Moores e.j.moores@aston.ac.uk College of Health and Life Sciences, Aston University, Birmingham, B4 7ET, UK
Supplemental data for this article can be accessed online at https://doi.org/10.1080/03075079.2023.2245424.

on-campus experience. Ewing (2021) reported an international study of vice-chancellors' views of COVID-19's impact on higher education and what the future might hold; some reported that COVID-19 has 'legitimised' online learning and most supported some kind of 'blended learning' in the future which involves aspects of online and on-campus teaching (see also Saha et al. 2022; Xie, Siau, and Nah 2020).

A UK survey of 33,726 students suggested that 45% wanted a mix of on campus and online learning, 42% wanted only on-campus provision, and the remainder (13%) wanted online learning (Jisc 2022). Clearly, some element of online learning is wanted by many universities and students, but the potential effects of such a pedagogical shift on student engagement, learning and attainment have not yet been fully elucidated. There is some evidence that blended learning can be highly effective (e.g. Bernard et al. 2014; Spanjers et al. 2015; Strelan, Osborn, and Palmer 2020) and can increase engagement and learning (e.g. Akçayir and Akçayir 2018; Haugom 2022; Northey et al. 2015; Snowball 2014), although that success may depend on the teaching methods employed (e.g. Foo, Cheung, and Chu 2021; Lin, Huang, and Lin 2022; Orlov et al. 2021) or on the student (e.g. Cacault et al. 2021; DePaola, Gioia, and Scoppa 2022).

There is a strong relationship between student engagement and attainment, even if that relationship is not considered causal (see e.g. Credé, Roch, and Kieszczynka 2010 for a review, or the meta-analysis by Lei, Cui, and Zhou 2018), and the same relationship appears to hold with online teaching. For example, Virtual Learning Environment (VLE) activity and attainment have been shown to correlate for both in-person and online courses (Agudo-Peregrina et al. 2014; Macfadyen and Dawson 2010; Mogus, Djurdjevic, and Suvak 2012; Morris, Finnegan, and Wu 2005; Ramos and Yudko 2008; Summers, Higson, and Moores 2021; Waheed et al. 2020; You 2016). In a longitudinal study of 106 students Saqr et al. (2023) found that whilst high engagement over a long period and several time points across a four-year programme of study can predict high achievement, a short period of disengagement can predict lower achievement in students with low engagement. Peach et al. (2019) similarly showed that whereas high performers showed 'diverse temporal engagement', low performers were more likely to fall in a massed learning or 'crammers' cluster, illustrating that the consistency as well as the amount of engagement is important. Many higher education providers now use learning analytics systems to track students' engagement with their courses, both online and in-person. One consequence of this is that universities have an opportunity to use these data to conduct interventions that support attainment and progression for students based on their engagement, rather than their demographics (Foster and Siddle 2020; Summers, Higson, and Moores 2021). This could enable institutions to use such systems to improve equality of outcomes between different student groups (Khalil, Slade, and Prinsloo 2023; Williamson and Kizilcec 2021). The use of learning analytics data to target such interventions has the advantage of avoiding stigmatisation of students with particular characteristics, as well as increasing the efficiency of the learning process. Celik et al. (2023) systematically reviewed publications on how learning analytics have been used during the pandemic, providing insight on usage, challenges, and potential benefits of this technology.

The impact of the COVID-19 pandemic on teaching in Higher Education has provided a useful 'natural experiment' to help evaluate the impact of different teaching methods. Orlov et al. (2021) conducted a study of economics students results pre- and peri-pandemic across four different institutions and argued that it was not 'who you teach', but 'how you teach' that mattered. Student characteristics showed little association with a general decline observed in student performance during the pandemic (see also Lin, Huang, and Lin 2022). Summers, Higson, and Moores (2023) provided an analysis of learning analytics data on engagement for different groups of students and found that students from disadvantaged backgrounds showed differences in digital engagement pre- vs. peri-pandemic, on average watching fewer recorded lectures peri-pandemic than they had previously; this might be expected to affect attainment, although this was not explicitly tested. Altindag, Feliz, and Tekin (2021) reported an increase in average grades following the in-person teaching/ online assessment regime, but more so for students living in areas with better broadband technology, suggesting that access to technology may play a role in student attainment, at least when assessment is online.

The time before and during the pandemic have provided a unique opportunity to compare the impact of different teaching and assessment approaches on students from disadvantaged and non-disadvantaged backgrounds. Using measures of attainment and engagement from a large and diverse sample the data available to us allowed further investigation of Orlov et al.'s contention that who you teach does not matter. This work expanded on Summers, Higson and Moores' (2023) findings of different engagement patterns amongst disadvantaged students pre- and peri-pandemic, to include examination of the relationship between engagement and attainment; Summers, Higson, and Moores (2023) did not analyse attainment data. Specifically, the effect of three teaching and assessment regimes on student attainment was investigated for both disadvantaged (IMD [Index of multiple deprivation] quintiles 1–2) and non-disadvantaged (IMD quintiles 3–5) students: (i) in-person teaching with in-person assessment, (ii) in-person teaching with online assessment and (iii) online teaching with online assessment. We also explored the extent to which differences in attainment were associated with different patterns of engagement and how this interacted with the mode of teaching and assessment.

Materials and methods

Sample participants

Aston University is a medium-sized UK university of around 11,000 students which has a student body that is more diverse, in terms of both ethnicity and socio-economic background, than most other UK universities. For this study, the records of three cohorts of first year full-time home undergraduate students who began their studies in 2018/19, 2019/20 and 2020/21 were analysed (initial sample of 5,182 students). First year undergraduates were used because the learning analytics system data was only available across all three years of interest for the first-year cohorts.

Students who were repeating a year, did not complete their first year (2018/19 and 2019/20 cohorts) or were not listed as current at the end of the academic year in June 2021 (2020/21 cohort) were removed from the sample. Additionally, and to enable normalisation of learning analytics data (see below), students were omitted from the sample if their course had not been monitored by the learning analytics system in each of the three years, as were students on courses with fewer than ten members. For the remaining students their home postcode was matched to a UK-wide adjusted index of multiple deprivation (IMD) quintile. IMD is a readily available measure of disadvantage which has a moderate correlation with permanent income deprivation (Jerrim 2021) and is therefore likely to be related to digital poverty, although this is a complex and arguably 'non-binary' issue (see e.g. Hernandez and Faith 2022). IMD quintiles are not normally comparable between the countries of the UK, but Abel, Barclay, and Payne (2016) derived an adjustment such that indices from three of the constituent countries of the UK can be compared with each other. This adjustment has been updated for the most recent 2020 indices by Parsons (2021) and was used here; namely that indices for Scotland, Wales and Northern Ireland were adjusted to be

Table 1. Demographics for each of the three cohorts comprising the sample for each (T)eaching and (A)ssessment mode.

		Disadvantaged			Non-disadvantaged			
Cohort year	Mode	Female	Male	Total	Female	Male	Total	Overall
2018/9	T:In person A:In person	441	334	775	306	337	643	1418
2019/0	T:In person A:Online	535	407	942	302	263	565	1507
2020/1	T:Online A:Online	603	465	1,068	305	284	589	1657

comparable to those from England. After removing students whose IMD quintile could not be identified, due to unrecognised postcodes, a sample of 4,582 students remained (see Table 1). The UK Office for Students (OfS) considers students from IMD quintiles 1 and 2 as meeting widening participation criteria (disadvantaged), whilst students from quintiles 3–5 are considered non-disadvantaged; students were divided into two categories that aligned with this distinction.

There were 29 courses amongst the sample of students covering a wide range of the courses offered, including Business, Engineering, English Language and Literature, Computer Science, Law, Mathematics, Optometry, Pharmacy and Psychology. Some courses, such as Optometry and Pharmacy, required limited on-campus attendance during online teaching to participate in laboratory classes.

Sample data

For each student, two items of attainment data were obtained; end of year mark and credits obtained. The end of year mark (%) is the average mark of all first-year modules that the student had completed. Each credit corresponds to ten hours of study, and a student is considered to have completed a stage of study (year) if they obtain 120 credits. A student can only reach 120 credits if they pass all their modules (score at least 40%) and can only progress to the following stage of study if they do so.

Engagement data from the university's learning analytics system were also analysed. All undergraduate modules are managed through a Virtual Learning Environment (VLE), where announcements, timetables, online synchronously delivered lectures, and course materials can be accessed. Since 2018, in-person attendance at lectures and seminars has been electronically recorded by students swiping their identity card, though neither attendance nor the act of recording attendance is compulsory for home students. Additionally, all lectures are recorded and available through the VLE via a lecture capture system (LCS). The learning analytics system aggregates the log data from the VLE, attendance recording system and lecture recordings on a daily basis. Following Summers, Higson, and Moores (2023), we used the data from three sources: (i) VLE course accesses: number of times the student accessed course materials, (ii) Attendance: total number of in-person classes and online classes that the student 'attended' – note that this could encompass either in-person attendance or virtual attendance, but the presentation was 'live' and 'synchronously' delivered in each case, and (iii) LCS: number of times the student viewed recorded lectures. A fourth data feed, number of library books checked out, was dropped from this analysis as library use was so low in 2020/21 that the median and modal number of books checked out was zero.

Analyses

All the statistical analyses were computed using R 4.2.0 (R Core Team 2022). Linear- and Logistic mixed-effects models were computed using *lmer* and *glmer*, respectively, from the package *lme4* (Bates et al. 2015). The significance of the effects of the main factors of the linear mixed-effects models were evaluated using the package *lmerTest* (Kuznetsova, Brockhoff, and Christensen 2017) which implemented the Satterthwaite approximation to estimate the denominator degrees of freedom of the F statistic. Estimated marginal means of the linear models were computed using *emmeans* (Lenth 2021).

The relationship between end of year mark, teaching mode and IMD was explored with the following linear mixed-effects model:

$$\begin{aligned} Mark_i = \quad & \beta_0 + \beta_1 Mode_i + \beta_2 IMD_i + \beta_3 Mode_i IMD_i + \\ & \beta_4 Sex_i + \beta_5 Ethnicity_i + b_i + \epsilon_i \end{aligned} \tag{1}$$

where $Mark_i$ is the mark of the ith student, β_0 is the intercept, β_x are the coefficients, *Mode* is the teaching mode which could be Teaching: In-person; Assessment: In-person, Teaching: In-person;

Assessment: Online, or Teaching: Online; Assessment: Online, and *IMD* is either Q12 (disadvantaged) or Q345 (non-disadvantaged). The controls are *Sex* (Female or Male) and *Ethnicity* (White, Asian, Black, Other or Unknown).[1] The term b_i is a zero-mean normally distributed parameter to account for any differences in mean grade between courses for each teaching/assessment mode. The term ϵ is the residual error. Note, the *anova* function was used to compare this model with two others and it was found to be the best fitting model by minimum Akaike information criterion (Akaike 1974); the other models considered had the random-effects structure of Course nested by IMD, and where the random-effects structure was just a simple random intercepts model for Course. Nesting by students did not occur since students only appeared in a single academic year (repeating students were excluded from analyses).

We also examined the effects of mode of teaching and assessment mode and IMD on pass (120 credits obtained) and fail (<120 credits obtained) with the following logistic model:

$$\log\left(\frac{P(Credits_i = 120)}{P(Credits_i < 120)}\right) = \begin{array}{r} \beta_0 + \beta_1 Mode_i + \beta_2 IMD_i + \beta_3 Mode_i IMD_i + \\ \beta_4 Sex_i + \beta_5 Ethnicity_i + b_i \end{array} \quad (2)$$

where $P(credits = 120)$ and $P(credits < 120)$ are the probabilities of students obtaining 120 credits, or not, respectively. We used this split of either achieving 120 credits or not, because all students who do not achieve 120 credits must do some form of extra work (either essays or resitting examinations) in order to continue their studies in the next year or stage, although we accept that those students with less than 120 credits may be a highly heterogenous group.

Clustering of engagement data

Cluster analysis (*kmeans* clustering in base R) was used to identify groups of students who were similar in terms of their learning analytics data. Prior to clustering, for each student, the average weekly data for the three sources of learning analytics data was computed for the 21 teaching weeks of the 2018/19, 2019/20 and 2020/21 academic years. For the 2018/19 and 2019/20 academic years, synchronous or live teaching was conducted entirely on campus, whereas for 2020/21 teaching was conducted almost entirely online, either synchronously (live) or asynchronously (e.g. pre-recorded lectures). For each student the weekly data – VLE course accesses, Attendance, LCS – was then normalised to z-scores by course and cohort; i.e. for a given data source in each course and in each year the mean and standard deviation were computed and each students' weekly data was normalised by subtracting the relevant mean and dividing by the relevant standard deviation; note that this normalised data was used only to generate the clusters by making the data comparable across all year groups, all other analyses were carried out on the raw data.

The optimum number of clusters for the whole dataset was assessed using gap analysis with *clusGap* from the *cluster* library (Maechler et al. 2022). The gap statistic is a measure of the intra-cluster sum of squares for a given clustering – its *compactness* – which can be compared with that from a reference dataset with no clustering. The *clusGap* function was used to generate reference datasets by bootstrapping ($n = 500$). Gap statistics were computed for up to six clusters and the *globalSEMax* rule (Dudoit and Fridlyand 2002) demonstrated that the maximum gap statistic occurred for three clusters (see Figure S1). For each cluster, the relationship between attainment (end of year mark), IMD and mode was explored using the linear model defined above.

Results

Average end of year marks

There was a significant effect of teaching mode on attainment [$F_{(2,\ 25.7)} = 3.887$, $p = .033$] with attainment being highest in the cohort for which teaching was in-person and assessment was online.

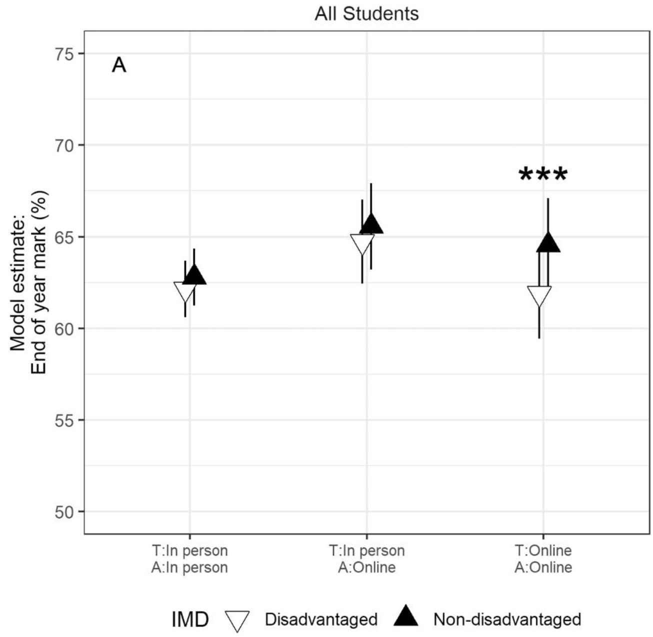

Figure 1. Estimated marginal means of the linear mixed-effects model for end-of-year mark (y axis), (T)eaching/(A)ssessment mode (x axis) and IMD (different symbols). Error bars indicate 95% confidence interval. Asterisks indicate significant differences in attainment between disadvantaged and non-disadvantaged students for the relevant teaching/assessment mode (* $p < .05$; ** $p < .01$; *** $p < .001$).

There was also a significant attainment gap [main effect, IMD, $F_{(1,\ 4525.1)} = 17.456$, $p < .001$]; disadvantaged students achieved lower marks than their non-disadvantaged peers. A significant interaction between teaching mode and IMD [$F_{(2,\ 4518.3)} = 4.145$, $p = .016$] showed that this difference was largest in the online teaching and assessment cohort (Figure 1). Post-hoc comparisons indicated no significant attainment gap between disadvantaged and non-disadvantaged students when teaching and assessment were in-person (0.6% pts [$t_{(4497.6)} = 1.145$, $p = .252$]) or when teaching was in-person and assessment online (0.8% pts [$t_{(4518.0)} = 1.509$, $p = .131$]). Only when teaching and assessment were both online, was there a significant attainment gap (2.6% pts [$t_{(4514.4)} = 4.893$, $p < .001$]). To put this in context, an attainment gap of 2.6% pts is equivalent to a quarter of a degree classification boundary and, if the gap were eliminated, would result in an uplift of the classification for 22% of disadvantaged students when teaching and assessment were online. It should be noted that the interaction is driven by both the non-disadvantaged students achieving higher marks in the online assessment regimes, as well as by the disadvantaged students achieving lower marks. Both Sex [$F_{(1,\ 4431.2)} = 70.978$, $p < .001$] and Ethnicity [$F_{(4,\ 4526.8)} = 24.990$, $p < .001$] were also significantly associated with attainment[2], with white students and female students obtaining higher marks. These are well-known attainment gaps for both characteristics (e.g. HEFCE 2018).

Credits obtained

The estimated probabilities from the logistic model of students passing the 1st year by obtaining 120 credits at the first attempt are plotted in Figure 2. The coefficients of the logistic model are provided in Table 2. To aid interpretation of the logistic model and determine the effects of IMD and teaching mode on credits, pairwise comparisons between disadvantaged and non-disadvantaged students were calculated for each teaching mode (see Table S1); only during online teaching and assessment was the log odds of obtaining 120 credits in disadvantaged students lower than that of their

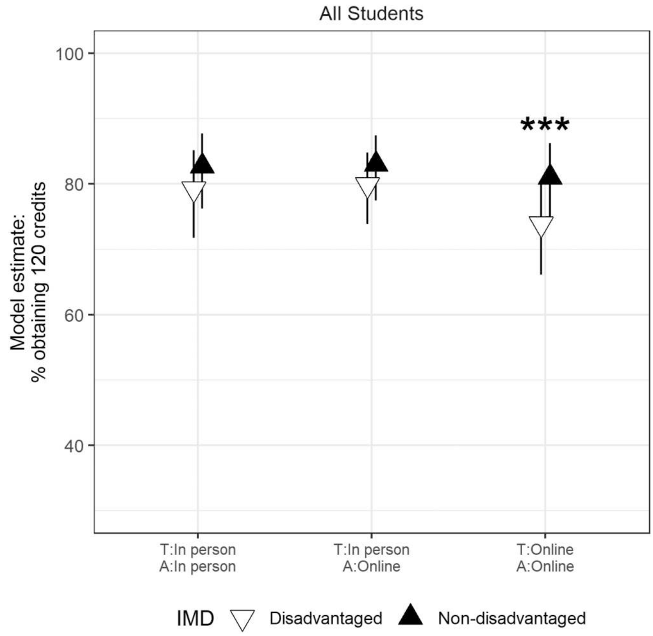

Figure 2. Estimated probabilities of students obtaining 120 credits (y axis) from a logistic mixed-effects model of the relationship between credits obtained (120 credits or <120 credits), (T)eaching/(A)ssessment mode (x axis) and IMD (different symbols). Error bars indicate 95% confidence interval. Asterisks indicate significant differences in credits obtained between disadvantaged and non-disadvantaged students for the relevant teaching/assessment mode (* $p < .05$; ** $p < .01$; *** $p < .001$).

counterparts (log odds = 0.412, $Z = 3.296$, $p < .001$). For the in-person teaching and assessment cohort and the in-person teaching/ online assessment cohort the differences were not significant (log odds 0.226, $Z = 1.641$, $p = .101$ and log odds 0.210, $Z = 1.565$, $p = .118$ respectively). The change in log odds between different teaching modes for disadvantaged and non-disadvantaged students are in Table 3. Only one comparison was significant, indicating a significant decrease in log odds of obtaining 120 credits for disadvantaged students in fully online vs. fully in-person teaching and assessment. To determine if there was an interaction between Mode and IMD we compared the difference in log odds between disadvantaged and non-disadvantaged students for the three different contrasts in teaching mode. None of these contrasts were significant (see Table S2), indicating there was no interaction between IMD and mode.

Table 2. Results of a logistic mixed-effects model between credits obtained (either 120 or <120), (T)eaching/(A)ssessment mode, and IMD (Disadvantaged or Non-disadvantaged).

Factor	Coefficient	SE	Z	p
(Intercept)	1.340	0.208	6.436	<.001
T:In person/A:Online	0.040	0.163	0.247	.805
T:Online/A:Online	−0.300	0.216	−1.388	.165
Non-disadvantaged	0.226	0.138	1.641	.101
SexMale	−0.427	0.080	−5.364	<.001
EthnicityAsian	−0.063	0.109	−0.574	.566
EthnicityBlack	−0.550	0.128	−4.312	<.001
EthnicityOther	−0.268	0.164	−1.639	.101
EthnicityUnknown/NA	−0.537	0.310	−1.730	.084
T:In person/A:Online:Non-disadvantaged	−0.016	0.190	−0.087	.931
T:Online/A:Online:Non-disadvantaged	0.185	0.183	1.013	.311

Coefficients are the βs from Equation (2); e.g. the line 'T:Online/A:Online' corresponds with β_1 when teaching and assessment are both online. The intercept is the estimate of the log odds for Disadvantaged students when Teaching and Assessment were in-person (β_0 in equation (2)). Other coefficients are changes in the log odds with reference to Disadvantaged students and Mode = T:In person/A:In person.

Table 3. Pairwise comparisons for the log odds of obtaining 120 credits between different (T)eaching/(A)ssessment modes for either disadvantaged students or for non-disadvantaged students.

Contrast	Estimate	SE	Z	p
Disadvantaged				
T:In person/A:In person – T:In person/A:Online	−0.040	0.163	−0.247	.805
T:In person/A:In person – T:Online/A:Online	0.300	0.216	1.388	.165
T:In person/A:Online – T:Online/A:Online	0.340	0.160	2.126	.033
Non-disadvantaged				
T:In person/A:In person – T:In person/A:Online	−0.024	0.185	−0.128	.898
T:In person/A:In person – T:Online/A:Online	0.114	0.232	0.491	.623
T:In person/A:Online – T:Online/A:Online	0.138	0.190	0.726	.468

Estimate is the modelled difference in log odds between obtaining 120 credits or not for the stated contrast.

As Sex and Ethnicity were simple covariates in the logistic model (see equation (2)) interpretation of their effects can be read directly from Table 2. Male students were significantly less like likely to achieve 120 credits than female students, and Black students were significantly less likely to achieve 120 credits than white students.[3]

Engagement and attainment during the different teaching modes

The cluster analysis used to identify groups of students who were similar in terms of their learning analytics data revealed three clusters that can be broadly defined as:

(1) High engagers: Students who had high attendance of synchronously delivered teaching (online synchronous and/or in-person), LCS views, and VLE usage,
(2) Synchronous-teaching engagers: Students with high attendance of synchronously delivered teaching (online synchronous and/or in-person), but low LCS views, and average VLE usage, and
(3) Low engagers: Students whose engagement was low for attendance of synchronously delivered teaching (online synchronous and/or in-person), LCS views, and VLE usage.

Summary engagement data for each of the three clusters is in Table 4 which reveals how interaction with the LCS and VLE increased as teaching moved online during the pandemic and, due to a reduction in the number of synchronous teaching sessions (whether in-person or online), attendance decreased. The differences in engagement between the clusters are also apparent, e.g. synchronous-teaching engagers and high engagers have similar attendance profiles but different LCS views and VLE engagement.

Table 4. Cluster summary statistics: Number of students (*n*), and mean and standard deviation of the weekly levels of Attendance, LCS views and VLE Course Accesses broken down by cluster and (T)eaching/(A)ssessment mode.

Mode	n	Attendance	LCS	VLE course access
High engagers				
T:In person & A:In person	249	4.859 (2.111)	2.698 (1.983)	59.207 (31.928)
T:In person & A:Online	264	5.603 (2.397)	3.848 (2.432)	52.468 (32.128)
T:Online & A:Online	392	4.031 (1.898)	9.453 (5.027)	106.078 (47.135)
Synchronous-teaching engagers				
T:In person & A:In person	584	5.043 (1.680)	0.681 (0.739)	36.116 (18.360)
T:In person & A:Online	655	5.658 (2.041)	1.103 (1.091)	30.154 (17.487)
T:Online & A:Online	516	3.894 (2.053)	4.560 (3.558)	66.283 (33.215)
Low engagers				
T:In person & A:In person	585	2.680 (1.517)	0.680 (0.861)	23.410 (14.355)
T:In person & A:Online	588	3.099 (1.697)	0.837 (1.081)	18.434 (13.517)
T:Online & A:Online	749	1.704 (1.292)	2.857 (2.939)	35.905 (22.905)

Note that clustering was performed on the complete dataset (normalised as described in the text) but is presented separately for each teaching mode.

Overall, about 20% of students were 'high engagers', 38% were 'synchronous-teaching engagers' (in-person and/or online) and 42% were 'low engagers', but there was variation in the proportion of disadvantaged and non-disadvantaged students in the different clusters across teaching/assessment modes, in particular with fewer disadvantaged students in the low engager cluster during in-person teaching and assessment (see Figure 3). To explore how the relationship between IMD and engagement varied with teaching mode, a series of three chi-squared analyses were performed. Only for the in-person teaching and assessment regime were engagement cluster and IMD significantly related ($\chi^2_{(2)} = 15.594$, $p < .001$), with disadvantaged students being more likely to be 'high engagers' and less likely to be 'low engagers', whilst the proportion of 'synchronous-teaching engagers' was similar for both groups. When teaching moved online, the proportion of students in the 'synchronous-teaching engagers' cluster reduced substantially for both groups (on average between 9 and 13% pts for non-disadvantaged and disadvantaged students respectively). Neither the online teaching and online assessment teaching regime nor the in-person teaching with online assessment regime showed significant associations between cluster and IMD ($\chi^2_{(2)} = 1.693$ and $\chi^2_{(2)} = 2.757$ respectively). In other words, there was a more comparable pattern of engagement between these groups under both online conditions.

We investigated whether attainment in each of the different teaching regimes was affected by engagement and disadvantage in a similar way to that found for all students (Figure 1). Figure 4 (A–C) shows the model-estimate end of year marks achieved by each IMD group under the three different teaching regimes and by engagement cluster. Figure 4(A–C) shows generally lower attainment for disadvantaged students within each cluster and each teaching regime, but most notably amongst the low engagers during online teaching and assessment. We ran three linear mixed-effects models, one for each teaching mode (see Table 5). The attainment gap between disadvantaged students and their peers was greatest in the low engagement cluster and when teaching

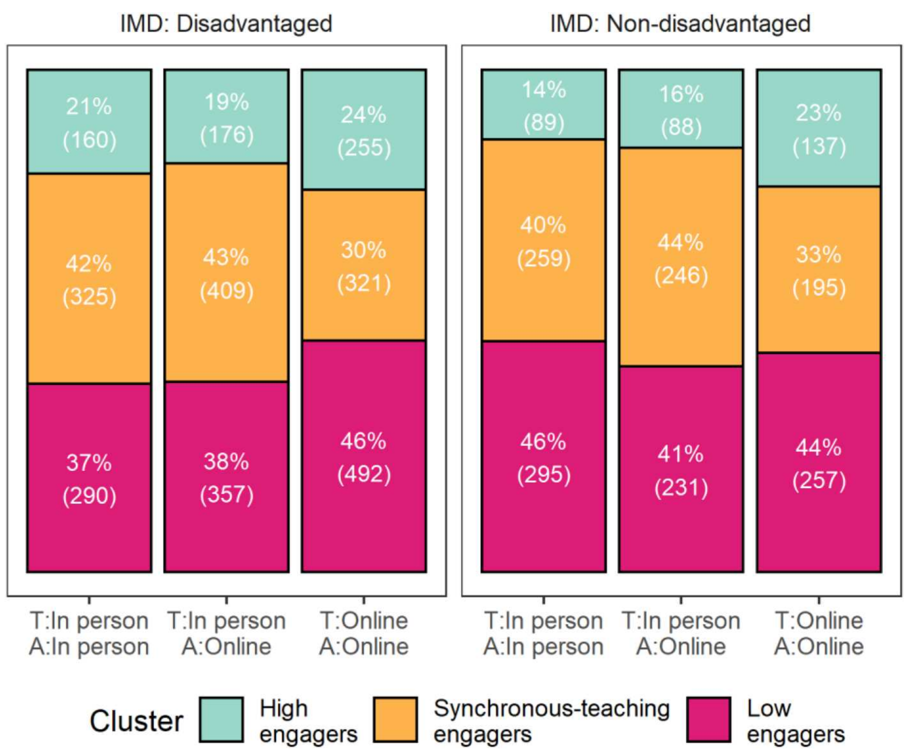

Figure 3. Relative proportions of students in each cluster by IMD (left panel = disadvantaged, right panel = non-disadvantaged) and (T)eaching/(A)ssessment mode (different bars). Numbers in brackets are the count of students in each cluster and IMD.

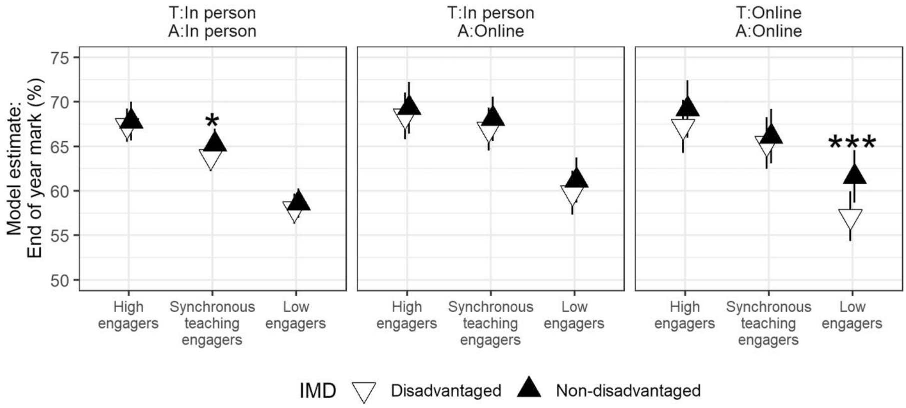

Figure 4. Estimated marginal means of the linear mixed-effects model for each (T)eaching/(A)ssessment mode (different panels), end-of-year mark (y axis), cluster (x axis) and IMD (different symbols). Error bars indicate 95% confidence interval of the mean. Asterisks indicate significant differences in attainment between disadvantaged and non-disadvantaged students for the relevant cluster (* $p < .05$; ** $p < .01$; *** $p < .001$).

Table 5. Results of ANOVA for three linear-mixed effects models, one for each teaching/assessment mode, relating end of year mark with Cluster and IMD.

Mode	Effect	*F*	*df*	*p*
T:In person/A:In person	Cluster	137.636	(2, 1385.5)	<.001
	IMD	2.722	(1, 1397.8)	.099
	Sex	0.880	(1, 1381.9)	.348
	Ethnicity	15.558	(4, 1395.5)	<.001
	Cluster:IMD	0.443	(2, 1386.1)	.642
T:In person/A:Online	Cluster	131.645	(2, 1468.9)	<.001
	IMD	5.105	(1, 1473.6)	.024
	Sex	13.931	(1, 1491.2)	<.001
	Ethnicity	11.804	(4, 1475.6)	<.001
	Cluster:IMD	0.081	(2, 1471.0)	.922
T:Online/A:Online	Cluster	83.740	(2, 1620.4)	<.001
	IMD	14.288	(1, 1627.1)	<.001
	Sex	20.918	(1, 1645.9)	<.001
	Ethnicity	6.841	(4, 1626.6)	<.001
	Cluster:IMD	3.970	(2, 1620.8)	.019

was online (4.5% pts, see Figure 4(C)). IMD was a significant factor when teaching was in-person and assessment was online ($F_{(1,1473.6)} = 5.105$, $p = .024$), and when teaching and assessment were both online ($F_{(1,1627.1)} = 14.288$, $p < .001$), but not when teaching and assessment were both in-person ($F_{(1,1397.8)} = 2.722$, $p = .099$). Only when teaching and assessment were online was there a significant interaction between cluster (engagement) and IMD ($F_{(2,1620.8)} = 3.970$, $p = .019$), with pairwise comparisons indicating the largest attainment gap in the low engager cluster (see flagged differences in Figure 4 and Table S3). That is, the main driver of the initial two-way interaction found between IMD and Teaching Mode for all students (Figure 1) was the attainment gap in the low engagers cluster in online teaching and assessment. During fully online teaching and assessment the non-disadvantaged low engagers appeared to maintain the benefits of online assessment shown by other groups, whilst the disadvantaged low engagers showed a decline; a cascade of disadvantage.

Discussion

In summary, our results showed that students from disadvantaged backgrounds (as defined above) fared worse than their peers during online teaching and assessment, both in terms of obtaining full credits and marks obtained (see also Bird, Castleman, and Lohner 2022; Kofoed et al. 2021). Whilst disadvantaged students were more likely than their peers to be in the high engagement cluster for in-person teaching and assessment, the association between disadvantage and engagement disappeared thereafter. However, despite more similar patterns of engagement in the two groups, the attainment gap was larger during online teaching and assessment, particularly for the group in the low engagement cluster. Disadvantaged students during the period of online teaching and assessment therefore showed a 'triple whammy' of effects; a greater likelihood to be in the lower engagement cluster than previously (in turn associated with lower attainment), lower overall attainment associated with their background, and lower attainment associated with the online teaching and assessment regime (at least in comparison to teaching in person/ assessment online conditions).

Our findings could be interpreted as a sign of digital poverty and, for some students, this may have been the case. Altindag, Feliz, and Tekin (2021) found a weaker increase in grades during in person teaching/ online assessment for those students living in areas with inferior broadband technology. Fifty-four percent of the disadvantaged students, however, remained in the high-engagers or synchronous-teaching engagers clusters during online teaching and assessment; clusters for which attainment remained high. A large proportion of students from both groups fell into the low engagement cluster, although it was for the disadvantaged students that this proportion increased most during online teaching, albeit only to a level similar to the non-disadvantaged group. Thus, the pattern of engagement alone cannot explain the attainment gap between these groups during online teaching and assessment.

Our results are in contrast with those of Orlov et al. (2021) which suggested that who you teach does not matter, and with El Said's (2021) results that found no negative effects of online learning, but align well with findings from several other studies (e.g. Bird, Castleman, and Lohner 2022; Kofoed et al. 2021; Rodríguez-Planas 2022; Summers, Higson, and Moores 2023) that reported some negative effects of online learning, particularly for students already disadvantaged in other ways. Of course, there are many things which changed during the pandemic unrelated to teaching and these may have also affected both engagement and attainment. For example, illness and financial insecurity were known to have affected disadvantaged groups more than their peers (see e.g. Rodríguez-Planas 2022) and studying at home is also likely to be more difficult for disadvantaged groups, whether or not there is a lockdown or appropriate digital infrastructure available (see, e.g. Bashir et al. 2021). Fabian et al. (2022, 5) coined the term *e-learning capital* as '*a measure of self-expressed ability and resources to utilise the online learning environment*' and found that it was positively correlated with study skills engagement during the pivot to online learning; whether students from disadvantaged backgrounds have lower e-learning capital and would thus be less likely to thrive during online teaching remains an open question.

Alternatively, the widening attainment gap during online teaching could be influenced by cohort differences. Students seeking entry to university did not sit examinations in summer 2020, meaning that – for the majority of students – teacher assessed grades (TAGs) were used for admissions purposes, rather than examinations nationally marked and moderated. In addition, much of the 2020/1 cohort would have been without their usual schooling (teaching or revision) for approximately six and a half months prior to university enrolment.

Although we identified three distinct patterns of student engagement, there was no cluster that showed lower engagement with synchronous-teaching and high digital engagement (LCS views and VLE accesses). Anecdotally, many students noted the benefits of synchronous sessions being online, specifically including the ability to anonymously ask questions in the 'chat' function of the software used (see also Hollister et al. 2022). There may be an advantage of attending synchronous teaching over pre-recorded content (see Fabriz, Mendzheritskaya, and Stehle 2021). Our division of participants into three clusters cannot be considered definitive. Although we used a well-established and widely used method of clustering, based on the best available criteria, cluster analysis can only ever be considered as exploratory, and variations in the agglomeration method used, or the variables included, may have produced somewhat different classifications. However, the clusters produced were intuitively plausible and provided understandable results.

Our findings build on those of Summers, Higson, and Moores (2023) that found changes in average engagement of disadvantaged students peri-pandemic but did not include any analysis of attainment in terms of either marks or credits obtained. In addition, analyses using clusters in the present study were able to show how students in each group moved between clusters during the different teaching regimes, precluding the possibility that average measures of engagement were affected by only a few students. In the present study, we also showed that the divergence in grade attainment between advantaged and disadvantaged students during the online teaching and assessment regime was driven principally by students in the low engagement cluster, further potential evidence of digital poverty and of the relationship between engagement and attainment.

The most significant limitation of this study is that the pandemic had profound effects on peoples' lives that extended well beyond teaching, and these may have been differentially affected different groups of students. It is also possible that the effects of enforced online teaching during a pandemic may not generalise to times when students have chosen online learning as their preferred medium of study (see e.g. Kizilcec, Makridis, and Sadowski 2021 for contrasting results). Similarly, these effects, found when staff had to hastily adapt to online teaching (see, e.g. Hodges et al. 2020; Orlov et al. 2021), may not apply in post-pandemic times now that staff are much more experienced in teaching in this format.

There are also some limitations related to the quality of the learning analytics data used in this study. For example, the learning analytics system measure of digital engagement only counted individual login events and did not capture other factors that may have may influenced students' experience, such as what type of device they are using (e.g. mobile phone, tablet, or PC), or whether the internet connection allowed them to watch, listen or contribute fully; the effects of digital poverty may therefore go beyond simply the ability to log in and be 'non-binary' (see Hernandez and Faith 2022).

Additionally, our analyses did not investigate any temporal aspects of engagement; there is increasing evidence that consistency of engagement as well as the total amount of engagement may be important (see e.g. Peach et al. 2019; Saqr et al. 2023; Summers, Higson, and Moores 2021). There are always some doubts over the validity of learning analytics data as it is typically incomplete, and our data is no exception (see e.g. Selwyn 2019 for a discussion). To overcome some of the potential issues of incomplete data, our data were normalised before clustering, to ensure that the student groups that emerged were driven by relative engagement within each teaching regimen and not by differences in the absolute levels of engagement between them.

Other limitations include the use of IMD as a measure of disadvantage. The IMD is an area-based measure that is an amalgamation of several indices of deprivation and, like all area-based measures IMD is subject to error in that i.e. non-deprived persons may live in disadvantaged areas and vice versa (Jerrim 2021), so the attainment gap due to IMD that we report here is likely to underestimate the real impact of deprivation. Similarly, as ethnicity was associated with IMD status in our sample, we may have underestimated the impact of disadvantage here by including ethnicity as a control variable. Analyses were conducted both with and without these controls and the attainment gap was indeed wider – and significant for all teaching modes – without them, this difference reflecting the influence of well-known ethnicity and gender attainment gaps. Finally, the results of this study are unlikely to generalise to countries that are less digitally developed (see, e.g. Adnan and Anwar 2020; Giatman, Siswati, and Basri 2020).

Although recent studies have reported on effects of the pandemic on engagement using learning analytics systems (e.g. Xu and Wilson 2021; Zhang, Taub, and Chen 2021), and on marks (e.g. Orlov et al. 2021), there has been a relative lack of research on equality issues using learning analytics (Williamson and Kizilcec 2021; but see Hlosta et al. 2021; Summers, Higson, and Moores 2023), despite the fact that learning analytics systems have valuable data to reveal insights about learners, learning, courses, and the university itself. Summers, Higson, and Moores (2023, 7) also found differential effects for asynchronously versus synchronously delivered digital material, suggesting '*a more nuanced, than a* *simplistic or all-encompassing view of "digital poverty"*'. Overall, therefore, it seems reasonable to speculate that different delivery and assessment methods may differentially affect different groups of students because of factors including – but not necessarily exclusive to – digital poverty. In other words, recorded online lectures as a sole or main source of provision, seems to reduce engagement in students, particularly those from disadvantaged backgrounds. In turn, lower engagement is associated with lower attainment. For providers wanting to close attainment gaps, this has important implications for their broader education strategy; 'how you teach' and 'who you teach' both matter.

It would be advisable for Higher Education providers to consider (and monitor) the potential implications of their education strategy on different groups of students in terms of engagement and attainment as well as reported preferences (see also Celik et al. 2023). The broader context of on-campus experience may offer something that the online experience does not (see e.g. Buhl-Wiggers, Kjærgaard, and Munk 2023). Providers wishing to provide hybrid or online learning could endeavour to monitor engagement with a view to providing interventions to ensure good engagement by all students, or experiment with different formats of online delivery to ensure that students show good engagement with their studies in a way that allows them to fulfil their potential. There is increasing evidence that patterns of engagement and disengagement – including at different times of the year – may be at least as important as average engagement across the year (e.g. Saqr et al. 2023; Summers, Higson, and Moores 2021). Further research should attempt to elucidate the elements which make an on campus or a blended experience successful.

Notes

1. A control for prior attainment using entry tariffs was not available due to the use of teacher-assessed grades in 2020 A-levels.
2. Note that in our sample 30% of white students are disadvantaged, whereas 66% of Asian students and 73% of Black students are disadvantaged, indicating an association between IMD and Ethnicity. We also conducted the analysis without inclusion of Sex and Ethnicity variables; this analysis showed that (i) the attainment gap was significant (wider), (ii) the attainment gap was present for all teaching modes, and (iii) there remained a significant interaction between teaching mode and IMD.
3. As above, we repeated the analysis without Sex and Ethnicity as covariates. In this case a significant main effect of IMD was found, but no effect of – or interaction with – teaching mode.

Acknowledgements

The data for this study were part of a larger dataset collected by the University for various functions including equal opportunities monitoring and enhancing the student experience. No new or additional data were collected and all data were anonymised. The analyses conducted were part of a project which monitors fair access and success. Student consent to their data being used for these purposes upon accepting their offer to study. RJS was supported by an award to EM from Aston University's Teaching Research Fund.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Data availability statement

Due to difficulties in properly anonymising the dataset we are unable to share the data associated with this article.

ORCID

Robert J. Summers http://orcid.org/0000-0003-4857-7354
Adrian P. Burgess http://orcid.org/0000-0002-0977-8105
Helen E. Higson http://orcid.org/0000-0003-3433-2823
Elisabeth Moores http://orcid.org/0000-0003-3997-0832

References

Abel, G. A., M. E. Barclay, and R. A. Payne. 2016. "Adjusted Indices of Multiple Deprivation to Enable Comparisons Within and Between Constituent Countries of the UK Including an Illustration Using Mortality Rates." *BMJ Open* 6: e012750. https://doi.org/10.1136/bmjopen-2016-012750.

Adnan, M., and K. Anwar. 2020. "Online Learning Amid the COVID-19 Pandemic: Students' Perspectives." *Journal of Pedagogical Sociology and Psychology* 2 (2): 45–51. https://doi.org/10.33902/JPSP.2020261309.

Agudo-Peregrina, A. F., S. Iglesias-Pradas, M. Conde-Gonzalez, and A. Hernandez-Garcia. 2014. "Can We Predict Success from Log Data in VLEs? Classification of Interactions for Learning Analytics and Their Relation with Performance in VLE-Supported F2F and Online Learning." *Computers in Human Behavior* 31: 542–50. https://doi.org/10.1016/j.chb.2013.05.031.

Akaike, H. 1974. "A New Look at the Statistical Model Identification." *IEEE Transactions on Automatic Control* 19 (6): 716–23. https://doi.org/10.1109/TAC.1974.1100705.

Akçayir, G., and M. Akçayir. 2018. "The Flipped Classroom: A Review of its Advantages and Challenges." *Computers & Education* 126: 334–45. https://doi.org/10.1016/j.compedu.2018.07.021.

Altindag, D. T., E. S. Feliz, and E. Tekin. 2021. *Is Online Education Working?* NBER Working Paper Series: 29113. https://doi.org/10.3386/w29113.

Bashir, A., S. Bashir, K. Rana, P. Lambert, and A. Vernallis. 2021. "Post-COVID-19 Adaptations; the Shifts Towards Online Learning, Hybrid Course Delivery and the Implications for Biosciences Courses in the Higher Education Setting." *Frontiers in Education* 6. https://doi.org/10.3389/feduc.2021.711619.

Bates, D., M. Maechler, B. Bolker, and S. Walker. 2015. "Fitting Linear Mixed-Effects Models Using lme4." *Journal of Statistical Software* 67: 1–48. https://doi.org/10.18637/jss.v067.i01.

Bernard, R. M., E. Borokhovski, R. F. Schmid, R. M. Tamin, and P. C. Abrami. 2014. "A Meta-Analysis of Blended Learning and Technology Use in Higher Education: From the General to the Applied." *Journal of Computing in Higher Education* 26 (1): 87–122. https://doi.org/10.1007/s12528-013-9077-3.

Bird, K. A., B. L. Castleman, and G. Lohner. 2022. "Negative Impacts from the Shift to Online Learning During the COVID-19 Crisis: Evidence from a Statewide Community College System." *AERA Open* 8. https://doi.org/10.1177/23328584221081220.

Buhl-Wiggers, J. A., A. Kjærgaard, and K. Munk. 2023. "A Scoping Review of Experimental Evidence on Face-to-Face Components of Blended Learning in Higher Education." *Studies in Higher Education* 48 (1): 151–73. https://doi.org/10.1080/03075079.2022.2123911.

Cacault, M. P., C. Hildebrand, J. Laurent-Lucchetti, and M. Pellizzari. 2021. "Distance Learning in Higher Education: Evidence from a Randomized Experiment." *Journal of the European Economic Association* 19 (4): 2322–72. https://doi.org/10.1093/jeea/jvaa060.

Celik, I., E. Gedrimiene, A. Silvola, and H. Muukkonen. 2023. "Response of Learning Analytics to the Online Education Challenges During Pandemic: Opportunities and key Examples in Higher Education." *Policy Futures in Education* 21 (4): 387–404. https://doi.org/10.1177/14782103221078401.

Credé, M., S. G. Roch, and U. M. Kieszczynka. 2010. "A Meta-Analytic Review of the Relationship of Class Attendance with Grades and Student Characteristics." *Review of Educational Research* 80 (2): 272–95. https://doi.org/10.3102/0034654310362998.

DePaola, M., F. Gioia, and V. Scoppa. 2022. "Online Teaching, Procrastination and Students' Achievement: Evidence from COVID-19 Induced Remote Learning." *SSRN Electronic Journal*, 15031. https://doi.org/10.2139/ssrn.4114561.

Dudoit, S., and J. Fridlyand. 2002. "A Prediction-Based Resampling Method for Estimating the Number of Clusters in a Dataset." *Genome Biology* 3 (7), https://doi.org/10.1186/gb-2002-3-7-research0036.

El Said, G. R. 2021. "How Did the COVID-19 Pandemic Affect Higher Education Learning Experience? An Empirical Investigation of Learners' Academic Performance at the University in a Developing Country." *Advances in Human-Computer Interaction*, https://doi.org/10.115/2021/6649525.

Ewing, L. A. 2021. "Rethinking Higher Education Post COVID-19." In *The Future of Service Post-COVID-19 Pandemic, Volume 1. The ICT and Evolution of Work*, edited by J. Lee and S. H. Han. Singapore: Springer. https://doi.org/10.1007/978-981-33-4126-5_3.

Fabian, K., S. Smith, E. Taylor-Smith, and D. Meharg. 2022. "Identifying Factors Influencing Study Skills Engagement and Participation for Online Learners in Higher Education During COVID-19." *British Journal of Educational Technology* 00: 1–22. https://doi.org/10.1111/bjet.13221.

Fabriz, S., J. Mendzheritskaya, and S. Stehle. 2021. "Impact of Synchronous and Asynchronous Settings of Online Teaching and Learning in Higher Education on Students' Learning Experience During COVID-19." *Frontiers in Psychology*, https://doi.org/10.3389/fpsyg.2021.733554.

Foo, C. C., B. Cheung, and K. M. Chu. 2021. "A Comparative Study Regarding Distance Learning and the Conventional Face-to-Face Approach Conducted Problem-Based Learning Tutorial During the COVID-19 Pandemic." *BMC Medical Education* 21 (1): 141. https://doi.org/10.1186/s12909-021-02575-1.

Foster, E., and R. Siddle. 2020. "The Effectiveness of Learning Analytics for Identifying At-Risk Students in Higher Education." *Assessment & Evaluation in Higher Education* 45 (6): 842–13. https://doi.org/10.1080/02602938.2019.1682118.

Giatman, M., S. Siswati, and I. Y. Basri. 2020. "Online Learning Quality Control in the Pandemic Covid-19 Era in Indonesia." *Journal of Nonformal Education* 6: 168–75. https://doi.org/10.15294/jne.v6i2.25594.

Haugom, E. 2022. "The Effect of Changing from Campus-Based to Digital Teaching on Student Attendance: A Case Study of Norwegian Business Students." *Heliyon* 8 (11): e11307. https://doi.org/10.1016/j.heliyon.2022.e11307.

HEFCE. 2018. *Differences in Student Outcomes: The Effect of Student Characteristics (ioe.ac.uk)*.

Hernandez, K., and B. Faith. 2022. *Measuring Digital Exclusion: Why Want is Counted is Also What Counts. Digit Data Commentary 01*. Digital Futures at Work Research Centre. https://digit-research.org/data_commentaries/measuring-digital-exclusion/.

Hlosta, M., C. Herodotou, M. Fernandez, and V. Bayer. 2021. "Impact of Predictive Learning Analytics on Course Awarding Gap of Disadvantaged Students in STEM." In: *22nd International Conference on Artificial Intelligence in Education, AIED* 2021, Lecture Notes in Artificial Intelligence. Springer.

Hodges, C., S. Moore, B. Lockee, T. Trust, and A. Bond. 2020. "The Difference Between Emergency Remote Teaching and Online Learning." *Educause Review* 27: 1–12. http://www.cetla.howard.edu/workshops/docs/The%20Difference%20Between%20Emergency%20Remote%20Teaching%20and%20Online%20Learning%20_%20EDUCAUSE%20(2).pdf.

Hollister, B., P. Nair, S. Hill-Lindsay, and L. Chukoskie. 2022. "Engagement in Online Learning: Student Attitudes and Behavior During COVID-19." *Frontiers in Education* 7. https://doi.org/10.3389/feduc.2022.851019.

Jerrim, J. 2021. *Measuring Disadvantage*. The Sutton Trust. Accessed 1 October 2022. https://www.suttontrust.com/wp-content/uploads/2021/05/Measuring-Disadvantage.pdf.

Jisc. 2022. *Student Digital Experience Insights Survey 2021/2: UK Higher Education (HE) Survey Findings*. Accessed 19 September 2022. https://www.jisc.ac.uk/reports/student-digital-experience-insights-survey-2020-21-uk-higher-education-findings.

Khalil, M., S. Slade, and P. Prinsloo. 2023. "Learning Analytics in Support of Inclusiveness and Disabled Students: A Systematic Review." *Journal of Computing in Higher Education*, https://doi.org/10.1007/s12528-023-09363-4.

Kizilcec, R. F., C. Makridis, and K. C. Sadowski. 2021. "Pandemic Response Policies' Democratizing Effects on Online Learning." *Proceedings of the National Academy of Sciences* 118 (11): 11. https://doi.org/10.1073/pnas.2026725118.

Kofoed, M., L. Gebhart, D. Gilmore, and R. Moschitto. 2021. "Zooming to Class? Experimental Evidence on College Students' Online Learning During Covid-19." *IZA Discussion Paper* 14356. https://doi.org/10.2139/ssrn.3846700.

Kuznetsova, A., P. B. Brockhoff, and R. H. B. Christensen. 2017. "lmerTest Package: Tests in Linear Mixed Effects Models." *Journal of Statistical Software* 82: 1–26. https://doi.org/10.18637/jss.v082.i13.

Lei, H., Y. Cui, and W. Zhou. 2018. "Relationships Between Student Engagement and Academic Achievement: A Meta-Analysis." *Social Behavior and Personality: An International Journal* 46: 517–28. https://doi.org/10.2224/sbp.7054.

Lenth, R. V. 2021. *emmeans: Estimated Marginal Means, aka Least-Squares Means*. R package version 1.6.2-1. https://CRAN.R-project.org/package=emmeans.

Lin, P. H., L. R. Huang, and S. H. Lin. 2022. "Why Teaching Innovation Matters: Evidence from a Pre- Versus Peri-COVID-19 Pandemic Comparison of Student Evaluation Data." *Frontiers in Psychology* 13. https://doi.org/10.3389/fpsyg.2022.963953.

Macfadyen, L. P., and S. Dawson. 2010. "Mining LMS Data to Develop an "Early Warning System" for Educators: A Proof of Concept." *Computers & Education* 54 (2): 588–99. https://doi.org/10.1016/j.compedu.2009.09.008.

Maechler, M., P. Rousseeuw, A. Struyf, M. Hubert, and K. Hornik. 2022. *cluster: Cluster Analysis Basics and Extensions*. R Package Version 2.1.4 – For New Features, see the 'Changelog' file (in the package source), https://CRAN.R-project.org/package=cluster.

Mogus, A. M., I. Djurdjevic, and N. Suvak. 2012. "The Impact of Student Activity in a Virtual Learning Environment on Their Final Mark." *Active Learning in Higher Education* 13 (3): 177–89. https://doi.org/10.1177/1469787412452985.

Morris, L. V., C. Finnegan, and S. S. Wu. 2005. "Tracking Student Behavior, Persistence, and Achievement in Online Courses." *The Internet and Higher Education* 8 (3): 221–31. https://doi.org/10.1016/j.iheduc.2005.06.009.

Northey, G., T. Bucic, M. Chylinski, and R. Govind. 2015. "Increasing Student Engagement Using Asynchronous Learning." *Journal of Marketing Education* 37 (3): 171–80. https://doi.org/10.1177/0273475315589814.

Orlov, G., D. McKee, J. Berry, A. Boyle, T. DiCiccio, T. Ransom, A. Rees-Jones, and J. Stoye. 2021. "Learning During the Covid-19 Pandemic: It is Not Who You Teach, But How You Teach." *Economics Letters*, 109812. https://doi.org/10.1016/j.econlet.2021.109812.

Parsons, A. 2021. *UK 2020 Composite Index of Multiple Deprivation*. https://github.com/mysociety/composite_uk_imd.

Peach, R. L., S. N. Yaliraki, D. Lefevre, and M. Balahona. 2019. "Data-Driven Unsupervised Clustering of Online Learner Behaviour." *NPJ Science of Learning* 4. https://doi.org/10.1038/s41539-019-0054-0.

Ramos, C., and E. Yudko. 2008. "Hits (Not "Discussion Posts") Predict Student Success in Online Courses: A Double Cross-Validation Study." *Computers and Education* 50 (4): 1174–82. https://doi.org/10.1016/j.compedu.2006.11.003.

R Core Team. 2022. *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing. https://www.R-project.org/.

Rodríguez-Planas, N. 2022. “Hitting Where it Hurts Most: COVID-19 and Low-Income Urban College Students.” *Economics of Education Review* 87. https://doi.org/10.1016/j.econedurev.2022.102233.

Saha, S. M., S. A. Pranty, M. J. Rana, M. J. Islam, and M. E. Hossain. 2022. “Teaching During a Pandemic: Do University Teachers Prefer Online Teaching?” *Heliyon* 8: e08663.

Saqr, M., S. López-Pernas, S. Helske, and S. Hrastinski. 2023. “The Longitudinal Association Between Engagement and Achievement Varies by Time, Students’ Profiles, and Achievement State: A Full Program Study.” *Computers & Education* 199: 104787. https://doi.org/10.1016/j.compedu.2023.104787.

Selwyn, N. 2019. “What’s the Problem with Learning Analytics?” *Journal of Learning Analytics* 6: 11–9. https://doi.org/10.18608/jla.2019.63.3.

Snowball, J. D. 2014. “Using Interactive Content and Online Activities to Accommodate Diversity in a Large First Year Class.” *Higher Education* 67 (6): 823–38. https://doi.org/10.1007/s10734-013-9708-7.

Spanjers, I. A. E., K. D. Könings, J. Leppink, D. M. L. Verstegen, N. de Jong, K. Czabanowska, and J. J. G. van Merriënboer. 2015. “The Promised Land of Blended Learning: Quizzes as a Moderator.” *Educational Research Review* 15: 59–74. https://doi.org/10.1016/j.edurev.2015.05.001.

Strelan, P., A. Osborn, and E. Palmer. 2020. “The Flipped Classroom: A Meta-Analysis of Effects on Student Performance Across Disciplines and Education Levels.” *Educational Research Review* 30: 100314. https://doi.org/10.1016/j.edurev.2020.100314.

Summers, R. J., H. E. Higson, and E. Moores. 2021. “Measures of Engagement in the First Three Weeks of Higher Education Predict Subsequent Activity and Attainment in First Year Undergraduate Students: A UK Case Study.” *Assessment & Evaluation in Higher Education* 46 (5): 821–36. https://doi.org/10.1080/02602938.2020.1822282.

Summers, R. J., H. E. Higson, and E. Moores. 2023. “The Impact of Disadvantage on Higher Education Engagement During Different Delivery Modes: A Pre- Versus Peri-Pandemic Comparison of Learning Analytics Data.” *Assessment and Evaluation in Higher Education* 48 (1): 56–66. https://doi.org/10.1080/02602938.2021.2024793.

Waheed, H., S. Hassan, N. R. Aljohani, J. Hardman, S. Alelyani, and R. Nawaz. 2020. “Predicting Academic Performance of Students from VLE Big Data Using Deep Learning Models.” *Computers in Human Behavior* 104: 106189. https://doi.org/10.1016/j.chb.2019.106189.

Williamson, K., and R. F. Kizilcec. 2021. Learning Analytics Dashboard Research Has Neglected Diversity, Equity, and Inclusion. In *Proceedings of the ACM Conference on Learning at Scale (L@S)*. Accessed 25 January 2023. https://rene.kizilcec.com/wp-content/uploads/2022/01/williamson2022review.pdf.

Xie, X., K. Siau, and F. F.-H. Nah. 2020. “COVID-19 Pandemic – Online Education in the New Normal and the Next Normal.” *Journal of Information Technology Case and Application Research* 22 (3): 175–87. https://doi.org/10.1080/15228053.2020.1824884.

Xu, Y., and K. Wilson. 2021. “Early Alert Systems During a Pandemic: A Simulation Study on the Impact of Concept Drift.” *LAK21: 11th International Learning Analytics and Knowledge Conference, April 2021*, 504–510. https://doi.org/10.1145/3448139.3448190.

You, J. W. 2016. “Identifying Significant Indicators Using LMS Data to Predict Course Achievement in Online Learning.” *Internet and Higher Education* 29: 23–30. https://doi.org/10.1016/j.iheduc.2015.11.003.

Zhang, T., M. Taub, and A. Chen. 2021. “Measuring the Impact of COVID-19 Induced Campus Closure on Student Self-Regulated Learning in Physics Online Learning Modules.” *LAK21: 11th International Learning Analytics and Knowledge Conference, April 2021*, 110–120. https://doi.org/10.1145/3448139.3448150.