

2021-02-10

Introducing immersive scheduling in a UK university: Potential implications for student attainment

Turner, Rebecca

<http://hdl.handle.net/10026.1/16919>

10.1080/0309877X.2021.1873252

Journal of Further and Higher Education

Taylor & Francis (Routledge)

All content in PEARL is protected by copyright law. Author manuscripts are made available in accordance with publisher policies. Please cite only the published version using the details provided on the item record or document. In the absence of an open licence (e.g. Creative Commons), permissions for further reuse of content should be sought from the publisher or author.

Introducing immersive scheduling in a UK university: Potential implications for student attainment

Dr Rebecca Turner*¹, Dr Oliver J. Webb¹ and Professor Debby R.E. Cotton²

¹Pedagogic Research Institute and Observatory, University of Plymouth, Plymouth, UK.

²Plymouth Marjon University, Derriford Road, Plymouth, PL6 8BH.

Email addresses: rebecca.turner@plymouth.ac.uk; oliver.webb@plymouth.ac.uk; dcotton@marjon.ac.uk

Orchid ID:

Dr Rebecca Turner <https://orcid.org/0000-0003-3953-618X>

Dr Oliver Webb <https://orcid.org/0000-0002-9412-157X>

Prof Debby Cotton orcid.org/0000-0001-7675-8211

*corresponding author

Twitter @DrBeckyTurner

Traditionally, undergraduates study several ‘long thin’ modules at the same time. Under ‘immersive scheduling’, students complete a ‘short fat’ module (i.e. a single subject studied over a compressed period), before moving onto other modules. This piece of social research capitalised on the introduction of immersive scheduling to the first year of undergraduate programmes at one UK university. Both semesters began with a short fat module, before students switched to studying long thin modules simultaneously. A novel within-subjects approach compared how individuals’ (N>3000) performed in immersively-delivered modules versus traditional modules. Overall, marks on immersively-delivered modules were significantly higher, with this

pattern replicated across semesters and in various demographic subgroups. This real-world evaluation complements existing between-subjects studies, where an identical module was delivered in immersive and traditional formats to separate cohorts. It offers further indication that immersive scheduling may be a beneficial pedagogic tool for enhancing student attainment.

Keywords: widening participation, attainment, inclusivity, first year induction

Introduction

Higher education institutions (HEIs) are experiencing a period of rapid and dramatic change (Barber et al., 2013). There has been an increasing focus on teaching in many countries, prompted by rising student fees and concurrent expectations of a quality experience. These factors have encouraged many HEIs to engage in curriculum transformations to enhance student learning, retention and attainment. One of the most frequent initiatives for transforming teaching is to re-structure the academic year, most commonly from semesters to terms (or vice versa). However, more changes can be made to the timetabling and delivery of programme – one such an approach is ‘immersive scheduling’.

Immersive scheduling (also known as ‘intensive mode teaching’ or ‘block scheduling’) denotes a situation where students are taught a single topic or module in concentrated bursts (e.g. a few weeks). This contrasts the traditional ‘long, thin’ model, where students simultaneously study multiple modules over several months. In both models, the time that students spend learning may be the same. However, proponents of immersive scheduling argue that it allows students to focus on and achieve mastery in particular elements of the curriculum, before moving on.

Some US universities introduced immersive scheduling in the 1990s in response to retention issues (Muraskin, 1998). The approach was treated with caution, amidst concerns that the learning experience may not be as effective or challenging as traditional delivery formats (Daniel, 2000; Traub, 1997). By 2005, however, 300 colleges and universities were delivering courses featuring immersive modules (Kucsera and Zimmero, 2010). Immersive scheduling has also been adopted in Australia, where flexible delivery models were used to attract part-time, mature or postgraduate learners (Burton and Nesbit, 2008; Davies, 2006).

Here immersive scheduling allowed students to balance study, work and family life (Burton and Nesbit, 2008). Interest in immersive scheduling is intensifying with an Australian university planning to adopt this approach wholesale, following promising results from pilot work (Times Higher Education, 2018). In order to inform contemporary policy making in HEIs, more insight is needed to the impact of immersive scheduling.

The requirement for further research is acute in the UK where, aside from limited application in several business schools (Davies, 2006), immersive scheduling has yet to be embraced. Research can inform institutional decision-making and convince students of what may be an unfamiliar approach. Students can harbour reservations prior to an immersive module (e.g. related to perceptions of workload and fatigue), though following completion, these concerns emerge as unfounded and students proactively choose immersive scheduling where available (Burton and Nesbit, 2008).

Studies record high levels of student satisfaction with immersive delivery (e.g. Burton and Nesbit, 2008; Daniel, 2000; Richmond et al., 2016; Scott, 2003), higher levels of motivation and enthusiasm; enhanced relationships with peers and course tutors; greater confidence in the subject studied; and an improved ability to balance the academic demands of the course (Kucsera and Zimmaro, 2010; Scott, 2003; Richmond et al., 2016). Recently, studies have started to consider the impact on attainment, often through analysis of ‘grades’ (e.g. percentage-based performance on a single module) which, despite being an imperfect proxy for learning, are ubiquitous in HE. Certainly, grades are used in ways that have significant impact on students’ prospects. In the UK, for example, attainment of a ‘good degree’ (typically, an average mark of $\geq 60\%$ across modules) can govern graduates’ access to postgraduate study or job opportunities (e.g. graduate training programmes). Student attainment represents a legitimate focus for investigation, with Burton and Nesbit (2008;

P.16) advocating, “*systematic comparison of students’ grades across traditional and block [immersive] formats*”. Current literature in this subdomain is limited in scale and scope - often reporting on student performance in a single immersive module. Nonetheless, results are promising. Studies have shown heightened performance (Richmond et al., 2016), with proportionally greater gains for students with a lower GPA (Karaksha et al., 2013). Consequently, there have been calls for further work that is both larger in scale and examines specific effects within different subgroups of learners (e.g. young versus mature students) (Davies, 2006; Kucsera and Zimmero, 2010).

The current study responds directly to these calls. It exploited a natural opportunity for studying student attainment in a UK university, following the introduction of immersive scheduling. In the 2015/16 academic year, a cross-institutional curriculum transformation sought to enhance transition to HE for first year undergraduates and increase retention. The central element was early immersion of students in their HE experience, via a re-structuring of year one in most academic programmes. Each semester began with an immersive ‘short fat’ module, after which students spent the rest of the semester studying simultaneous ‘long thin’ modules.¹ Importantly, this structure dictated that first year students experienced both module formats, without self-selecting delivery method.

Previous studies in this field followed ‘between-subjects’ designs, under which an identical module is delivered in immersive and traditional formats to entirely separate cohorts, whose performance is then compared. The primary benefit is that, aside from delivery format, core elements of the learning experience can be kept constant between the two conditions (i.e. form(s) of assessment, curriculum content and identify of the instructor).²

¹ For ease of reading, the two module formats will be described from here as ‘immersive’ and ‘traditional’.

² Although it should be noted that previous between-subjects studies do not describe any formal ‘fidelity’ measures i.e. systematic checks that delivery was consistent in all aspects other than delivery method.

There are, however, practical and methodological challenges with between-subjects approaches. In the UK, university ethics committees or the broad student body may not countenance full randomisation of individuals between immersive and traditional versions of a module. Studying even a single module in immersive format is likely to impact the scheduling of students' remaining modules, introducing potential inequity with the cohort who studied exclusively through traditional modules. Application of staunch scientific methods to participants in a real world context is unlikely to be approved (Dunn, 2008), especially given the increasing identification of some students as 'consumers' of HE (Tomlinson, 2017). Moreover, it may be difficult to uphold a core principle of randomisation - participants' 'blindness' to the fact that another cohort is receiving an alternative learning experience. As Karanicolas et al. (2010: 346) note, "*If participants are not blinded, knowledge of group assignment may affect their behaviour in the trial*". In the absence of randomisation, participants in between-subjects studies may have opportunity to choose a delivery format. Students' preferences for immersive delivery are entangled with other characteristics (e.g. confidence in their academic discipline; Burton and Nesbit, 2008), which could confound the relationship between delivery format and performance. Furthermore, without the laborious process of generating matched samples, it can be difficult to account for the effects of personal characteristics (e.g. self-efficacy, motivational outlook, prior subject knowledge, academic ability) which can be highly variable between individuals. If, overall, individuals in one cohort have more pronounced characteristics (e.g. high self-efficacy) versus their counterparts in the other group, the true impact of the experimental manipulation (i.e. delivery format) may be obscured. Finally, the requirement to run multiple versions of the same module is resource-intensive. Consequently, between-subjects studies are likely to be limited in terms of the number of students, and the range of modules/disciplines, over which effects of delivery format can be observed.

The curriculum transformation invited the use of alternative, ‘within-subjects’ methods, for exploring the association between delivery format and attainment. Individuals’ performance in immersively-delivered modules was compared against their grades from the traditionally-delivered modules. Without running multiple versions of the same modules, there was a naturally occurring opportunity to examine data from over 3000 students, spread across academic departments. A valid criticism of this within-subjects approach is that, alongside the delivery format, the immersive and traditional modules under comparison differed in terms of additional factors (i.e. assessment methods, curriculum content and instructors). Theoretically such uncontrolled elements might influence student grades, giving the spurious impression that delivery format is associated with attainment. However, through additional analyses and careful interpretation, we were able to judge the risk posed by these uncontrolled factors, and whether they fundamentally undermined any association between delivery format and attainment. The confounding influence of personal characteristics (e.g. motivational outlook) was largely negated; the same individuals completed immersive *and* traditional modules, such that these variables stayed constant. Given that between-subjects and within-subjects approaches have merits and imperfections, it is advantageous to use both methods to triangulate evidence around immersive scheduling. To summarise, this study sought to address the following research questions:

- What is the pattern of student attainment between modules delivered using immersive and traditional formats?
- Does the relationship between module format and attainment differ according to students’ social and/or demographic characteristics?

Methods

The host institution was a public university in southern England. Student records were used to identify undergraduates who were in the first year during the 2015/16 academic year - the point when immersive modules were introduced across the institution. The sample included students from 14 out of 16 departments that operated undergraduate programmes. The only groups not represented were students from the Schools of Medicine and Dentistry, where the atypical programme structure prevented the use of immersive modules. To be included in the sample, students needed to have completed at least one immersive module and one traditional module during a semester. An anonymised file was created featuring individuals' assorted first year module marks (%)³ and six demographic variables: gender (male versus female); age (<21 years versus \geq 21 years); disability (registered versus not registered); nationality (UK versus non-UK); entry tariff (quartile)⁴; and social class (groups I-VII)⁵.

The programmes in question followed a 'semesterised' structure comprised of two 15-week semesters (see Figure 1). This involved completion of 60 credits during the first half of the academic year (semester one) and a further 60 credits in the second half (semester two). Whilst some UK universities continue to operate modules that span an entire academic year, reports indicate that the majority of institutions take a semesterised approach (Times Higher Education, 2002). At the host university, each tranche of 60 credits normally comprised three 20-credit modules. In each semesters, the first four weeks were devoted to immersive delivery of a single 20-credit module (Figure 1). In the first semester, all students on a given

³ Data were excluded from modules that followed an atypical approach to assessment, which did not generate percentage marks (i.e. modules assessed on a 'pass/fail' basis only). Individuals' marks were also identified as outliers and excluded using the conventional boundaries of 1.5 x the interquartile range (See Tukey, 1977).

⁴ This refers to students' level of academic attainment at prior to entry into university. Under the framework, qualifications, such as A-levels, are converted into points to allow course providers to compare applications during the admissions process <https://www.ucas.com/file/63541/download?token=uz826-Cb>

⁵ <https://www.ons.gov.uk/methodology/classificationsandstandards/otherclassifications/thenationalstatisticsocioeconomicclassificationnssecrebasedonsoc2010>

programme completed the same immersive module. Collectively, these will be referred to as ‘Immersive 1’ or ‘I1’ Modules (n=47). In semester two, students on some programmes again completed a homogenous immersive module, whereas counterparts on other programmes were able to choose their immersive module from the broad portfolio offered across Schools. A total of 48 immersive modules were delivered in semester two, which will be collectively referred to as ‘Immersive 2’ or ‘I2’ Modules. Importantly, all I1 and I2 modules were created anew for the curriculum transformation. Therefore, comparison could not be made against traditional iterations of the same modules from previous years. After completing the respective I1 or I2 module, students spent the remainder of that semester in traditional modules. The traditional modules were taken simultaneously over an 11-week period; modules were taught for nine weeks and followed immediately by a two-week assessment period (Figure 1). A minority of the traditional modules were delivered on multiple occasions throughout the year, making it difficult to identify the semester in which individual students had participated. After exclusion of these, there were 152 traditional or T1 modules delivered in semester one, and 221 traditional or T2 modules delivered in semester two. For each student, marks from their respective T1 and T2 modules were reduced to two averages: $T1^{MEAN}$ and $T2^{MEAN}$. For clarity, the host university, in common with many UK HEIs, operates the following system for classifying undergraduate degrees: <40%= ‘fail’; 40-49%= ‘third-class degree’; 50-59%= ‘lower second-class degree’; 60-69%= ‘upper second-class degree’; $\geq 70\%$ = ‘first-class degree’.

Analysis of student attainment was carried out through two separate sets of mixed-design ANOVAs; one set for each semester. Under this approach, performance in immersive modules was compared against traditional modules from the same semester, making it possible to capture and contain any overarching influence of semester on performance. In all ANOVAs, the ‘within-subjects’ factor was delivery format i.e. I1 versus $T1^{MEAN}$ for analyses

of semester one, and I2 versus T2^{MEAN} for analyses of semester two. The ‘between-subjects’ factor was one of the demographic variables. Where demographic variables had more than two levels (i.e. UCAS entry tariff, social class), Tukey post hoc tests were employed. In all analyses p values of $< .05$ were used.

The design brief for I1 and I2 Modules stipulated that principles of inclusivity be followed and that all assessment be made via coursework.⁶ Assessment via coursework is perceived as fairer for students whose performance may be affected by a mental or physical health condition, as it allows individuals to manage the time and energy dedicated to the submission (Richardson 2015). Also coursework usually circumvents the need for modified assessment amongst certain student groups (e.g. extended exam time for dyslexic students). By contrast, T1 and T2 Modules featured assessment via coursework, unseen examination, practical observation, or a combination of methods. Consequently, additional analyses were needed, to probe if any difference in attainment between delivery types could be attributed to the methods of assessment; specifically, the ubiquitous use of coursework to assess immersive modules. To help unravel these factors T1 and T2 Modules, which all shared the same traditional delivery format, were pooled ($n= 373$) and arranged in four groups according to the proportion of assessment made via coursework. Comparison of average marks across these groups would identify relationships between assessment method and student performance.

The host university employs an anonymous marking policy applicable to coursework and unseen examinations. This means that the identities of students who submit coursework are only known to the marker(s) in the minority of cases where it is unavoidable (e.g. coursework

⁶ In this context, coursework represented an end of module assessment completed in a format other than an unseen exam. Examples include an essay, report, reflective account or creative output.

involving students presenting or performing a live task). Therefore marker bias should not be more pronounced in modules that make extensive use of coursework.

Results

In total 3141 students completed an I1 Module and one or more T1 Module, whilst 2628 completed an I2 Module and one or more T2 Modules. The reduced size of the latter sample – some 16% smaller - reflects student attrition. Many withdrawing prior to the completion of semester two modules re-registered for study the next year – as indicated by a ‘non-continuation’ rate of 7% (Higher Education Statistics Agency).⁷ A small number of individuals had marks for semester two modules but not semester one modules; this is likely to result from a delayed start (e.g. through illness), or where students were repeating selected modules as a consequence of failure in the preceding year. For analyses involving certain demographic variables, sample sizes were further reduced (see Figure 2). This was an anticipated result of missing demographic data. For example, data on social class and entry tariff is rarely available for non-UK students. Entry tariff can also be unavailable for students with atypical application profiles (e.g. those given credit on account of prior experiential learning).

Figure 3 shows average student performance in pooled T1 and T2 modules, stratified by the proportion of assessment via coursework. Whilst there was no compelling relationship, the highest performance (62%) was actually recorded in modules with the least coursework. According to these findings, it is unlikely that any heightened performance in immersive modules could be attributed to their exclusive use of coursework for assessment.

⁷ <https://www.hesa.ac.uk/data-and-analysis/performance-indicators/non-continuation>

Delivery Format

Across students, mean marks were 67.0% for I1 modules versus 64.7% for I2 modules. $T1^{MEAN}$ and $T2^{MEAN}$ were, respectively, 63.5% and 61.7%. Average performance on I1 and I2 modules was clearly several percentage points higher than traditional modules completed in the corresponding semester. At first glance, performance in I2 modules appears suppressed relative to I1 modules. Looking at the standardised effect sizes, however, the margin by which immersive modules outperformed traditional modules was very consistent in the two semesters (semester one $d = 0.27$, semester two $d = 0.26$).⁸ Thus, it appears that the dip in students marks for I2 modules versus I1 modules principally reflects an overall trend for student marks to be lower in the second semester.⁹ According to common interpretations, these effect sizes reflect a small but meaningful difference in student performance between immersive and traditional module formats (Cohen, 1988).

The impression of higher student performance in immersive modules was confirmed by statistically significant main effects for delivery format in every ANOVA analysis. Given that all analyses showed a consistent main effect of delivery format, related statistical output is only reported once, in the gender analysis below, so as to conserve space. Using the conventional option of box charts, results for different subgroups could not be plotted legibly on the same axes. Consequently, Figure 2 provides line charts instead.

⁸ Standardised effect sizes for repeated measures data computed according to advice from Morris and DeShon (2002).

⁹ A series of additional mixed-design ANOVAs were performed to explicitly compare student performance in I1 modules versus I2 modules, and $T1^{MEAN}$ versus $T2^{MEAN}$. These confirmed that performance in both immersive and traditional modules was significantly lower in semester two, compared to the corresponding type of module in semester one.

Gender

The analyses of both semesters showed a significant main effect for delivery format, with better performance in immersive modules relative to traditional modules (semester one, $F[1,3139]= 235.64, p < .001$; semester two, $F[1,2626]= 179.04, p < .001$). Both analyses also showed a significant main effect of gender, whereby females' overall performance exceeded that of males (semester one, $F[1,3139]= 28.77, p < .001$; semester two, $F[1,2626]= 37.44, p < .001$). Importantly, both analyses found a significant interaction between gender and delivery format, which is illustrated in Figures 2 (I) and 2 (II) (I1 versus T1^{MEAN}, $F[1,3139]= 8.38, p < .01$; I2 versus T2^{MEAN}, $F[1,2626]= 7.47, p < .01$). Whilst both genders showed heightened performance in immersive modules relative to traditional modules, the margin of enhancement was greater for males. Thus, the gap in performance between genders was narrowed during immersive modules, although there remained a statistically significant difference between males and females.

Age

The analysis of semester two data showed no significant main effects or interactions regarding age: performance was consistent between younger and older students (Figure 2 [IV]). In contrast, the semester one analysis showed a significant main effect of age, such that performance was, overall, higher in older students than their younger counterparts ($F[1,3139]= 6.97, p < .01$). Furthermore, there was a significant interaction between age and delivery format ($F[1,3139]= 53.43, p < .001$). As illustrated in Figure 2 (III), during the I1 modules there was no significant difference in performance between age groups. However, in the T1 modules that followed there was a clear deterioration in marks for younger students, contrasted with stable performance amongst older students. The pattern observed here for

older students is unique within the study. It represents the only instance where a subgroup's performance on the immersive module was not significantly better than the traditional modules in the same semester. To illustrate this irregularity, Figure 4 isolates individuals who recorded marks for all four time points and stratifies them according to demographic subgroup. With the exception of older students, all subgroups show a broadly consistent trajectory of performance: a downward trend across the year, with a spike in performance during each immersive module.¹⁰ Conversely, older students – denoted by the solid black line – showed strong and stable performance across the first three time points (I1, T1^{MEAN} and I2),¹¹ before a marked decline (T2^{MEAN}). This finding suggests that alongside the overarching difference in performance between semesters, which was mentioned earlier, certain subgroups may exhibit particular chronological patterns of performance.

Disability

Both analyses showed a small but significant main effect of disability whereby performance was, overall, lower amongst disabled students than their non-disabled peers (semester one, $F[1,3139]= 4.223, p < .05$; semester two, $F[1,2626]= 4.122, p < .05$). A lack of any significant interaction in semester one data contrasted with a small but significant interaction between disability and delivery format in semester two. As can be seen from Figure 2 (VI), the spike in performance in I2 modules, relative to T2 modules, was slightly enhanced in disabled students, versus their non-disabled peers.

¹⁰ The subgroups based on other demographics (e.g. entry tariff) also show this general trajectory but are not plotted in order to maintain legibility of the figure.

¹¹ Additional analyses confirmed no significant difference in performance of older students between I1 modules, T1^{MEAN} and I2 modules.

Entry Tariff as a Measure of Prior Achievement

In semester one, there was a significant main effect of entry tariff ($F[3,2606]= 29.98$, $p < .001$) but no interaction between entry tariff and delivery format. Figure 2 (VII) illustrates this main effect. It meant that overall performance was not significantly different between the first (lowest tariff) and second quartiles, or the third and fourth quartile, but performance in quartiles one and two was significantly lower compared with quartiles three and four.

Semester two data showed both a main effect of entry tariff ($F[3,2365]= 40.90$, $p < .001$) and a significant interaction between entry tariff and delivery format ($F[3,2365]= 8.70$, $p < .001$).

Close interrogation of these results is informative. As illustrated in Figure 2 (VIII), during I2 modules the relative performance of students in the respective tariff quartiles was consistent with findings for semester one. In the T2 modules that followed, however, the gap in performance between these subgroups broadened. $T2^{\text{MEAN}}$ data showed that performance amongst students in quartile one was now significantly lower than quartile two, with both these groups also significantly lower than quartiles three and four. Moreover, the difference between quartiles three and four now approached significance ($p= .054$). Those with lower entry tariffs (who already perform lower than their counterparts), appeared at particular risk of additional slippage during this latter part of semester two.

Nationality and Social Class

Analyses of these remaining factors showed a consistent pattern. There were main effects of each demographic variable on performance but no interactions with delivery format. In terms of nationality, performance amongst UK students was, overall, significantly higher than their non-UK counterparts (semester one, 65.5% versus 61.4%, $F[1,3139]= 33.22$, $p < .001$; semester two, 63.4% versus 60.5%, $F[1,2626]= 15.47$, $p < .001$). Regarding

social class, post hoc tests were conducted to examine in detail the main effects (semester one, $F[6,2560]= 3.95, p <.01$; semester two, $F[6,2116]= 2.69, p <.05$). Whilst these were not wholly conclusive, in both semesters the chief difference in performance appeared to be between those in class I (parents with higher managerial, administrative and professional occupations) and other students.

Discussion

Analyses indicated that student performance in immersive modules was, overall, higher than in traditionally-delivered modules from the same semester. Alongside these positive general findings, there was evidence of differential effects in men and women. Whilst both sexes performed better in immersive modules, compared with those delivered using a traditional format, the benefit was larger for men, thus narrowing the gap in attainment. This finding is welcome in light of a trend for lower male performance in HE, which has been in train since the 1990s (Richardson, 2015). The latest literature challenges a historic notion that particular forms of assessment disproportionately benefit male or female students (Hillman and Robinson, 2016; Richardson, 2015). It seems unlikely, therefore, that the heightened success observed amongst males during immersive modules was driven by the exclusive use of coursework for assessment. In terms of an alternative explanation, some research indicates that male students have less effective study behaviours than females (Webb & Cotton, 2018). With only a single module to focus on, the immersive approach may facilitate students in better organising their workload and approach to study. These concerns were highlighted in recent work relating to the first year experience, in which students were found to underestimate the amount of time required for HE study (e.g. Gibney et al. 2011) or the level at which they were expected to work (e.g. Maunder et al., 2013). Immersive scheduling could mitigate these issues and support students to develop essential self-regulation skills. To explain the general efficacy of immersive scheduling, previous authors

have indeed pointed towards its capacity for reducing the cognitive load faced by students, through limiting the volume of competing demands (e.g. Jansen, 2003; Richmond et al., 2015).

Results for age also indicate distinct patterns of responsiveness between subgroups. During semester one, older students showed no significant elevation in marks for the immersive module relative to traditional modules. This was the only instance where the performance of a subgroup did not appear enhanced under the immersive format. By contrast, younger students showed the prototypical profile of elevated performance in the first immersive module relative to other semester one modules. In the second semester, results for both subgroups were uniform; elevated performance in the opening immersive module was followed by deteriorating performance in the ensuing traditional modules. This pattern of effects is useful for considering additional explanations for the efficacy of immersive scheduling. With regard to cognitive load, there is no established age effect which explains why, in semester one, only younger students appeared to benefit from a rationalised schedule of learning. Instead, the results for age may reflect an additional social mechanism through which immersive delivery impacts performance. During immersive modules, participants were frequently re-exposed to the same classmates. Previous qualitative evaluation of immersive scheduling suggests that this helps facilitate the formation of peer learning communities (Turner et al. 2017) which have, in turn, been linked with heightened attainment (Buch and Spaulding, 2011). Social integration is certainly seen as having protective effects in terms of the more frequently researched domain of student drop-out (e.g. Tinto, 1994; Bean and Metzner, 1985; Gibney et al., 2011; Maunder et al., 2013; Trautwein and Bosse, 2017). In contrast, older students approach university study with higher levels of intrinsic motivation (Murphy and Roopchand, 2003) and less interest in socialisation, and whilst they

are interested in non-timetabled peer activities focussed on the curriculum, they may be unable to engage owing to competing commitments such as commuting and caring duties (Kasworm, 2014). Thus, the failure of the first immersive module to raise performance significantly amongst older students may reflect their strong motivation on entering HE; ambivalence towards extracurricular socialisation; or barriers to accessing curriculum-focussed peer activities. The interplay between delivery format, performance and age is clearly highly nuanced, however, underscoring the importance of further research – including qualitative studies - amongst different subgroups.

For three other demographic variables – nationality, disability, and social class – there were main effects such that students from the UK, those without a registered disability, and those from the highest category of social class, recorded higher marks than the corresponding subgroups. These patterns reflect previous findings (Morrison et al., 2005; Crawford, 2014; Richardson, 2010), which supports the credibility of the current data set. For nationality and social class there were no interactions with delivery format and, in the case of disability, a modest interaction in the second semester only. Broadly, this means that the enhancement in marks seen during immersive modules was consistent for UK versus non-UK students, disabled versus non-disabled students, and students from across the social class categories. Whilst immersive delivery did not compellingly narrow the attainment gaps reported for these demographic groupings, it did not worsen existing differences in performance. This finding was significant, given unease amongst some colleagues at the host institution that immersive delivery may disadvantage specific subgroups. It was, for example, important to allay concerns that a shorter window of delivery would deny non-UK students and dyslexic students (who likely account for a substantial proportion of those with a registered disability) sufficient time to engage with written course materials. Regrettably, other core demographics, such as ethnicity, were not accessible as part of the current investigation. There is an

imperative to include such factors in future studies of immersive scheduling, given evidence of an attainment gap between students with different ethnic backgrounds (Webb & Cotton, 2018).

In terms of prior educational attainment, previous studies were able to consider students' average university performance leading up to their immersive experience. For example, in the study by Karaksha et al. (2013) the immersive delivery occurred in year two of the undergraduate programme, such that students' first year GPA could be calculated to serve as a baseline. Whilst students who selected the immersive variant of the module tended to have a lower GPA, there was no significant difference in final marks between delivery formats, offering tentative evidence that immersive delivery helped narrow the range of performance between students (Karaksha et al., 2013). Such analysis was not possible in the current study where students' university experience opened with an immersive module, so divesting the chance to collect baseline data. By way of a proxy, students' entry tariff was available, which enabled insightful analyses. Over the first three time points (i.e. I1, T1^{MEAN}, and I2), the relative pattern of performance was maintained between students with different tariff profiles. By contrast, at the fourth and final time point performance became more divergent, with particular deterioration in marks for students with lower entry tariffs. It is clear that immersive delivery did not remove apparent attainment gaps between students, relating to their tariff profile. Nonetheless, it is possible that immersive delivery served here to moderate or delay a trend for weakening performance that is pronounced in lower entry students. Unfortunately, there is not at present any research documenting patterns of attainment between different student subgroups, at different stages in the university journey. Such 'event history analyses' have, however, been conducted in relation to student drop-out (Mabel and Britton, 2018). These confirmed that as students move deeper into a university

programme, the risk of drop-out for those with lower pre-entry performance rises disproportionately. Whilst student performance in the current study was tracked for a whole academic year - further than in previous studies from the area – there is a case for even lengthier follow-up. Longitudinal studies might address the question of whether immersive scheduling, whilst appearing to benefit performance in the immediate term, reduces students' opportunities to connect concepts and material from different modules in real time. This could manifest in impaired performance in capstone assessments such as final year dissertations.

Alongside evidence regarding the efficacy of immersive delivery, the current study produced a tangential finding of value to HE practitioners. Irrespective of delivery format, student performance was modestly but significantly lower in the second semester versus the first. In the UK, the majority of HEIs have chosen to organise modules into two discrete semesters, as opposed to the once traditional three-term model, under which material was studied across the whole academic year (Times Higher Education, 2002). Whilst authors have documented various motivations for adopting 'semesterisation' (Morris, 2000), there appears to be little evidence regarding the impact on student performance. The negative trend observed between semesters may reflect academic fatigue, although further qualitative research would be valuable in unpicking this phenomenon. If replicated, there is a wider implication of this finding for HE researchers. Investigations that measure the impact of interventions on student performance over lengthy periods may need to consider and account for this 'semester effect' statistically.

As noted, the real-world nature of the current study meant that it did not feature immersive and traditional iterations of the same module. Rather, the immersive and traditional modules under comparison not only differed in delivery format, but also the instructor, assessment methods and curriculum. Consequently, it was important to consider if

these factors could account for the patterns of student attainment that were observed. Starting with assessment, a sub analysis identified no association between module performance and assessment method, so discounting this potential confounding influence. This lack of association contrasts with some evidence from the 1990s, pointing to a positive relationship between marks and the use of coursework for assessment (Gibbs and Lucas, 1997). However, sophisticated analyses, which controlled for a range of demographic and contextual confounders (e.g. class size), indicated that the greatest difference lies between modules at the extremes of the spectrum, where 100% of assessment is made via coursework or examination (Simonite, 2003). When Simonite compared modules with 100% coursework against modules with *some* degree of coursework, there was no significant difference in marks for the majority of the disciplines studied (10 out of 13). Of nearly 500 modules (immersive and traditional) featured in the current study, the great majority fall into these two categories, with a mere 35 assessed by 100% examination.

Turning to other uncontrolled factors, the lecturers who worked on immersive modules could, in theory, have been more effective in terms of instruction, or more lenient with marking, compared with counterparts who delivered the traditional modules. A sense of scale is valuable here. The study draws from, respectively, 95 and 373 different immersive and traditional modules, developed by separate and autonomous teaching teams. Across this number of modules it is highly unlikely that more accommodating lecturers were accidentally allocated to those with immersive delivery. In addition, it is unlikely that lecturing staff massaged student grades in systematic fashion, to give a favourable impression of immersive delivery. Consistent with existing literature on staff attitudes to large curriculum transformations (e.g. Hasanefendic, Birkholz, North and van der Sija, 2017), lecturers in the current university were cautious, and in some cases opposed, to the introduction of immersive scheduling. Interestingly, previous large-scale analyses of pedagogical initiatives (e.g. active

learning) have found that identical courses, delivered by different instructors, showed little variation in attainment (Freeman et al., 2014). Thus, it appears that the impact of different teaching personnel on student grades may be overstated.

Aside from the identity of the instructor, certain immersive modules did incorporate specific features, which may have aided students who had suboptimal study behaviours. For example, some delivery teams issued structured timetables indicating how students might gainfully use time outside of class. Others featured co-delivery with learning developers, to embed study skills formally within the module. Both these approaches are supported in the literature (Thomas, 2012; Turner et al., 2017). Such initiatives were, however, sporadic and featured primarily in the immersive modules from semester one. These approaches cannot account for the heightened performance observed in immersive modules from the second semester. After a process of elimination, the most compelling explanation for the differences observed in student grades, remains the delivery format of modules.

In conclusion, the current within-subjects study corroborated findings from earlier between-subjects research. In a very large sample, which spanned multiple academic disciplines and demographic subgroups, student grades were significantly higher in immersive modules compared with those delivered in a traditional format. Results for certain groups (e.g. males) are particularly noteworthy, as immersive delivery appeared to narrow entrenched attainment gaps. The current findings underscore the need for further research, including qualitative approaches, to isolate the potential causal mechanisms for the association between delivery format and student attainment (e.g. reduced cognitive load, enhanced opportunities for social integration).

Disclosure statement: The authors declare that they have no competing interests.

References

Barber, M., K. Donnelly, S. Rizvi, and L. Summers 2013. *An avalanche is coming. Higher Education and the revolution ahead*. Institute for Public Policy Research.

Bean, J.P. and B.S. Metzner. 1985. "A conceptual model of non-traditional undergraduate student attrition." *Review of Educational Research* 55 (4): 485–

540.

Buch, K., and S. Spaulding. 2011. "The impact of a psychology learning community on academic success, retention, and student learning outcomes." *Teaching of Psychology* 38 (2): 71-77.

Burton, S. and P.L. Nesbit. 2008. "Block or traditional? An analysis of student choice of teaching format." *Journal of Management and Organisation* 14 (1): 4-19.

Cohen, J. 1988. *Statistical Power Analysis for the Behavioural Sciences*. Hillsdale: Lawrence Erlbaum.

Authors et al. 2016.

Crawford, C. 2014. *Socio-economic differences in university outcomes in the UK: drop-out, degree completion and degree class* (No. W14/31). IFS Working Papers.

Daniel, E. L. 2000. "A review of time-shortened courses across disciplines." *College Student Journal* 34 (2): 298–308.

Davies, M. 2006. "Intensive teaching formats: a review." *Issues in Education Research* 16 (1): 1-21.

Dunn, D. S. 2008. "Another view: In defence of vigour over rigor in classroom demonstrations." *Teaching of Psychology* 35 (4): 349-352.

Freeman, S., Eddy, S. L., McDonough, M., Smith, M. K., Okoroafor, N., Jordt, H., and Wenderoth, M. P. (2014). Active learning increases student performance in science, engineering, and mathematics. *Proceedings of the National Academy of Sciences*, 111 (23): 8410-8415.

Gibbs, G., and L. Lucas. 1997. "Coursework assessment, class size and student performance: 1984-94." *Journal of Further and Higher Education* 21 (2): 183-192.

Gibney, A., Moore, N., Murphy, F. and O'Sullivan, S. 2011. "The first semester of university life; 'will I be able to manage it all?'. *Higher Education* 62 (3): 351-366.

Hasanefendic, S., Birkholz, J.M., North, H and van der Sija, P. (2017). "Individuals in action: bringing about innovation in higher education." *European Journal of Higher Education*, 7(2): 101-119.

Hillman, N., and N. Robinson. 2016. *Boys to Men: The underachievement of young men in higher education-and how to start tackling it*. Oxford: Higher Education Policy Institute.

Karaksha, A., S. Anoopkumar-Dukie, G. Grant, A.K. Davey, N.S. Nirthanan, D. Arora, D. Hope, N. Bernaitis, A. McFarland, S. Hall, A. Haywood, C. Holani, L. Grant, R. Chess-Williams, and C. McDermott, C. 2013. "*Benefits of Intensive Mode Teaching to Improve Student Performance*". In I. O. Committee (Ed.). Presented at the ICERI 2013, Seville, Spain

Kasworm, C. E. 2014. "Paradoxical understandings regarding adult undergraduate persistence." *The Journal of Continuing Higher Education* 62 (2): 67-77.

Karanicolas, P. J., Farrokhyar, F., & Bhandari, M. (2010). Blinding: Who, what, when, why, how? *Canadian Journal of Surgery* 53 (5): 345-348.

Kucsera, J.V. and D.M. Zimmaro. 2010. "Comparing the effectiveness of intensive and traditional courses." *College Teaching* 58 (2): 62-68.

Lyall, C., L. Meagher, J. Bandola and Kettle, A. 2016. *Interdisciplinary provision in higher education: current and future challenges*. York: Higher Education Academy.

Mabel, Z., and T.A. Britton. 2018. "Leaving late: Understanding the extent and predictors of college late departure." *Social Science Research* 69 (1): 34-51.

Maunder, R.E., Cunliffe, M., Galvin, J., Mjali, S. and Rogers, J. 2013. "Listening to student voices: student researchers exploring undergraduate experiences of university transition." *Higher Education* 66 (1): 139-152.

Morris, H. 2000. "The Origins, Forms and Effects of Modularisation and Semesterisation in Ten UK - Based Business Schools." *Higher Education Quarterly* 54 (3): 239-258.

Morris, S. B., and R.P. DeShon. 2002. "Combining effect size estimates in meta-analysis with repeated measures and independent-groups designs." *Psychological Methods* 7 (1): 105-125.

Morrison, J., B. Merrick, S. Higgs, and J. Le Métails. 2005. Researching the performance of international students in the UK. *Studies in Higher Education*, 30 (3), 327-337.

Murphy, H., and N. Roopchand. 2003. "Intrinsic motivation and self-esteem in traditional and mature students at a post-1992 university in the north-east of England." *Educational Studies* 29 (2-3): 243-259.

Muraskin, L.D. 1998. *A structured freshman year for at-risk students*. US Department of Education, Washing, DC.

Richardson, J. T. 2010. "Course completion and attainment in disabled students taking courses with the Open University UK". *Open Learning*, 25 (2): 81-94.

- Richardson, J. T. 2015. "Coursework versus examinations in end-of-module assessment: a literature review." *Assessment & Evaluation in Higher Education* 40 (3): 439-455.
- Richmond, A. S., B.C. Murphy, L.S. Curl and K.A. Broussard. 2015. "The effect of immersion scheduling on academic performance and students' ratings of instructors." *Teaching of Psychology* 42 (1): 26-33.
- Scott, P. A. 2003. "Attributes of high-quality intensive courses." *New Directions for Adult and Continuing Education* 97 (1): 29-38.
- Simonite, V. 2003. "The impact of coursework on degree classifications and the performance of individual students." *Assessment & Evaluation in Higher Education* 28 (5): 459-470.
- Thomas, L. 2012. *Building Student Engagement and Belonging in Higher Education at a Time of Change: Final Report from The What Works? Student Retention & Success Programme*. London: Paul Hamlyn Foundation.
- Times Higher Education. 2002. "Students choke on bite-sized learning". Times Higher Education, October 4.
- Times Higher Education. 2018. "'Block teaching' model ripe for copycats". Times Higher Education, June 4.
- Tinto, V. 1994. *Leaving college: Rethinking the causes and cures of student attrition*. Chicago, IL: University of Chicago Press.
- Tomlinson, M. (2017). Student perceptions of themselves as 'consumers' of higher education. *British Journal of Sociology of Education* 38 (4): 450-467.
- Traub, J. 1997. "Drive-thru U: Higher education for people who mean business." *New Yorker*, October, 114-123.
- Trautwein, C. and Bosse, E. 2017. "The first year in higher education-critical requirements from the student perspective." *Higher Education* 73 (3): 371-387.
- Tukey, J. W. 1977. *Exploratory data analysis* (Vol. 2). Reading, MA: Addison-Wesley

Turner, R., Morrison, D., Cotton, D., Child, S., Stevens, S., Nash, P. & Kneale, P. (2017).

Easing the transition of first year undergraduates through immersive induction modules.

Teaching in Higher Education, 22(7): 805-821.

Webb OJ, Cotton DRE. (2018). Deciphering the sophomore slump: changes to student

perceptions during the undergraduate journey. *Higher Education*, 77, 173-190.