

A large-scale test of the relationship between procrastination and performance using learning analytics

Sophie H. Cormack^a, Laurence A. Eagle^b and Mark S. Davies^c

^aSchool of Psychology, University of Sunderland, Sunderland, UK; ^bCentre for Enhancement of Learning and Teaching, University of Sunderland, Sunderland, UK¹;

^cFaculty of Health Sciences and Wellbeing, University of Sunderland, Sunderland, UK

Published online Jan 7th 2020 (out later in print) in *Assessment and Evaluation in Higher Education*

Doi: 10.1080/02602938.2019.1705244

¹ Current affiliation: Head of Quality, Standards & Enhancement, Faculty of Education, Health and Wellbeing, University of Wolverhampton

A large-scale test of the relationship between procrastination and performance using learning analytics

Many studies have found a relationship between students' self-reported procrastination and their grades. Few studies have used learning analytic data as a behavioural measure of procrastination in order to predict performance, and there is no systematic research on how this relationship may differ across assessments or disciplines. In this study we analyse nine years' worth of institutional electronic submission records, a total of 73,608 assignment submissions, to examine the relationship between submission time and grades across assignments, students, courses, and disciplines in higher education. A significant negative relationship was found overall, with students who submitted closer to the deadline obtaining lower grades, however the size of the relationship was negligible, accounting for less than 1% of the variance in grades. The relationship varied significantly depending on student, assignment, course, and discipline.

Keywords: procrastination, learning analytics, assessment, academic performance

Aims and objectives

The aim of this paper was to examine whether digital records of higher education students' coursework assignment submission times could usefully be added to the analytic measures used to identify university students at risk of failure or drop-out. Previous studies are scarce, but several studies using coursework assessment completion time as a proxy for procrastination have found that later completion of assessments correlated with lower grades. These previous studies are limited to one or two courses in a small range of disciplines, so it is difficult to know to what extent it would be predictive for all students in all disciplines. This study examines the full range of

disciplines offered at a university, to investigate the relationship between submission time and grades, and whether it is consistent across students, assignments, courses, and disciplines. If so, this could be useful data to add to university-wide predictive analytics systems, but if the relationship is weak or context-dependent, its usefulness would be more limited.

Introduction

Learning analytics is a growing field, defined as ‘the measurement, collection, analysis and reporting of data about learners and their contexts, for purposes of understanding and optimising learning and the environments in which it occurs’ (description provided for conference website by Siemens, 2010). One of the most common uses of learning analytics is to identify students at risk of failure or drop-out, with the aim of improving retention (Dawson et al. 2014; Agudo-Peregrina et al. 2014). Typically this is done by looking at students’ frequency of access and cumulative time spent in their university’s virtual learning environment (VLE), interactions with other students and staff online, attendance, library use, marks for formative or summative assessments, and demographic data (e.g. De Freitas et al. 2014; Colvin et al. 2015). Relatively few studies have looked not just at what students do, but when they do it. Later engagement may indicate dilatory behaviour due to poorer self-regulation strategies or lower motivation, a pattern of behaviour subjectively experienced as procrastination, ‘the voluntary delay of an intended and necessary and/or [personally] important activity, despite expecting potential negative consequences that outweigh the positive consequences of the delay’ (Klingsieck 2013, 26).

Most university students procrastinate, by their own account, with 20-50% suffering chronic procrastination which they would rather overcome (Solomon and Rothblum 1984; Steel 2007). Procrastination in students has been linked with stress (Tice and Baumeister 1997), depression and anxiety (Solomon and Rothblum 1984; Rozental et al. 2015). A range of theories have attempted to explain procrastination (Klingsieck 2013), ranging from those linking procrastination with personality traits, to motivational approaches explaining procrastination as a gap between intention and action, owing to lack of intrinsic motivation or poor self-regulation. Situational approaches incorporate contextual factors, suggesting that procrastination occurs when the task to be done is particularly boring, difficult, pointless, or requires personal sacrifices. So some theories assume procrastination is an individual trait, which would predict that students with a tendency to procrastinate would show similar behaviour across all their higher education studies, while situational theories would predict that students procrastinate more on some tasks than others.

Procrastination has also been associated with lower grades. A meta-analysis by Kim and Seo (2015) found a significant weighted mean correlation of $r = -0.13$ between procrastination and grades, across different procrastination questionnaires, countries, age-groups, and assessment types. The size of the correlation was substantially larger when the performance measure was coursework grades ($r = -0.64$) rather than examination grades. The relationship between procrastination and performance was strongest for younger (compulsory sector) students. The strength of the relationship also varied depending whether procrastination was measured by self-report questionnaires (as in 68 of the studies reported in the 33 papers) or by external observation (14 studies), and whether performance was self-reported (20 studies) or

externally-assessed (62 studies). They found the strongest relationship ($r = -0.39$) between externally-assessed procrastination and externally-assessed performance. Kim and Seo argued that self-report is inaccurate and unreliable as students tend to exaggerate both their procrastination and their grades, so they favoured use of observational measures over self-report questionnaires. Of the papers that used external observation of procrastination to predict grades, five (Steel, Brothan, and Wambach 2001; Moon and Illingworth 2005; Howell et al. 2006; Rotenstein, Davis, and Tatum, 2009; Hensley 2014) used behavioural measures such as observed delay in completion of online quizzes, or assignment submission time. These behavioural measures are learning analytic measures which can be captured through digital traces.

A variety of learning analytic measures have been treated as proxy measures of procrastination, for example time of first activity in the VLE (e.g. McElroy and Lubich 2013; Baker et al. 2015), pace of VLE accesses to indicate cramming or spaced learning (e.g. Thille et al. 2014; Jo et al. 2016), and assignment submission time - how close to the deadline students completed assignments. Our study focuses on the last of these measures. Using submission time as a behavioural measure of procrastination assumes that students who hand in work close to the deadline have been putting off working on their assignments, leading them to submit very close to the deadline. Researchers have found that students spend increasing amounts of time on an assignment as the deadline approaches, and the same pattern has been observed in the timing of assignment submissions across a cohort, with the number of submissions increasing steeply as the deadline approaches, resembling a hyperbolic curve (Dewitte and Schouwenburg 2002; König and Kleinman 2005; Howell et al. 2006).

While there may be other reasons for later submission, such as making optimal use of information, illness, or leaving more time for proof-reading, it is fairly certain that students who submit further in advance of the deadline are less likely to be procrastinators. The advantage of submission time as a metric is that it is easy to extract from learning analytic data, so would be a good measure to include in calculations of at-risk students if it proves to be an important predictor of performance.

The earliest study that examined the relationship between submission time and performance was on American psychology students as they completed a series of online quizzes (Steel, Brothan, and Wambach 2001). Submission time had a moderate negative correlation with examination grade and a high negative correlation with coursework grade. These correlations were much higher than those between students' self-assessed level of procrastination and grades, suggesting that learning analytic measures of procrastination may be a better predictor of performance than questionnaire measures.

Moon and Illingworth (2005) carried out a similar study, also on psychology students in the USA, with similar findings. Levy and Ramim (2012) also looked at procrastination in online computer-marked quizzes, in information science students in the USA. They found that students who procrastinated until the day of the deadline performed significantly worse than those who completed the quizzes earlier. Hensley (2014) found, for anatomy students in the USA, a weak but significant negative correlation between quiz submission time and grades. Rotenstein, Davis and Tatum (2009) studied MBA students in the USA after they completed computer-marked homework assignments, and likewise found a weak but significant negative correlation

with homework grade. In the UK, Arnott and Dacko (2014) found that marketing students who submitted an assignment early obtained significantly more distinctions than later submitters. In Spain, Cerezo et al. (2016) found a significant negative relationship between psychology students' e-training assignment submission times and performance. In contrast, two studies of assessed discussion forum postings found no significant correlation between posting time and grades, in Canada (Howell et al. 2006) and Israel (Gafni and Geri 2010).

Overall, seven of nine studies found a significant relationship between submission time and performance, though the strength of relationships varied and only a small sub-set of all disciplines was tested. Gašević et al. (2016) pointed out the dangers of ignoring educational context, since learning analytics predictive of success for students in one discipline were not in another. Contextual factors that might have influenced the results of the nine studies on submission time and performance are largely unknown, so it would be valuable to determine to what extent the relationship varies depending on context, for example discipline.

Research questions

The main research question is to determine whether there is a reliable association between assignment submission time and grades. We predicted that students who submit closer to the deadline (a proxy measure of procrastination) would obtain lower marks than those who submitted well in advance of the deadline.

If a relationship was found, the second research question is to ask whether the relationship between submission time and grades would differ markedly from one student to another, or whether the costs of delay would be similar across all students.

Some (e.g. Chu and Choi 2005; Choi and Moran 2009; Corkin, Yu, and Lindt 2011; Kim, Fernandez, and Terrier 2017) have suggested that procrastinating students fall into two types, with only 'passive' procrastinators showing poorer outcomes. Students who are confident and delay as a deliberate strategy, so-called 'active procrastinators', thrive on pressure. So some students may not display a negative relationship between delay and grades.

The third research question is whether or not the relationship varies across assignments, as procrastination may be more costly for some assignments than others. Finally, we asked whether the relationship might vary in different courses (modules) or disciplines.

Methods

Data

The data for this study were all electronic assignment submissions for an entire university over a 9-year period (April 2006 to April 2015). The university is a 'post-1992' university located in the North of England, UK. Not all assignments were submitted electronically, so the data is not a complete representation of the university's assignment submissions. The file obtained from the electronic submission database contained 385,668 records each representing an individual student's assignment submission.

Data cleaning and sampling

The data were prepared for statistical analysis in the following ways. All records with no marks associated with them (73% of the sample) were discarded, as irrelevant for

answering our research questions. This could for example have included unmarked drafts, formative assessments, or late work which was not graded.

All records where students submitted after the deadline were discarded. These cases are likely to have included students who were offered extensions due to disabilities or extenuating circumstances, and possibly some students given a late work penalty. Submissions were also excluded if submitted by staff on behalf of a student.

Any assignments where all students submitted within a short timescale of less than 6 hours were excluded (these could have been exam submissions, or assignments that staff had forgotten to open to students until very late). As deadlines in the system were often recorded as one minute before the hour (e.g. 15:59), one minute was added to all the deadlines to ensure that students who submitted on time were not recorded as late. Some deadlines were recorded as a few minutes after the hour (e.g. 10:02, 14:10) suggesting staff allowed a small leeway: these were not altered, apart from adding the extra minute.

Most submissions were marked on the percentage scale, with a pass mark of 40%, however a few of the assignments had been marked using a different denominator (e.g. out of 20), and, as long as the marks awarded were consistent with this setting, those assignment marks were converted to the percentage scale. In some cases it was not clear whether assignments were marked on the full percentage scale, so as a precaution, we also excluded any assignments with only one submission awarded 50% or less, or any assignments with multiple submissions where the median mark was less than 40% and the highest mark 50% or less. Any individual submissions given a mark

over 100 were also excluded. Any assignments where students were only awarded only two or three different marks (e.g. all students received either 0% or 40%) were discarded as representing pass/fail assessments rather than those graded on a scale.

Finally, outliers were removed following the assumption that assignment submissions follow a hyperbolic curve with very few students submitting early, and increasingly large numbers of students submitting as the deadline approaches (Dewitte and Schouwenburg 2002; König and Kleinman 2005; Howell et al. 2006). Under this assumption, outliers at the 'early' end of the range would be expected, and therefore have not been removed. Outliers at the 'late' end of the scale, falling substantially after the bulk of submissions, were removed because this would not fit the expected curve, and would suggest that the deadline was reset to encompass a few students who were permitted to submit at a later date, after the bulk of the cohort. A submission was defined as an outlier if it was submitted more than 1.5 interquartile ranges after the later quartile. Quartiles were calculated using the interpolation method using Python's *numpy* package (Oliphant 2006).

At the end of this cleaning and sampling process, 73,608 records remained for analysis, 71% of the marked submissions, representing 19% of the original sample.

Variables

The following variables were utilised in this study:

The explanatory variable (predictor) was submission time. A measure of how close to the deadline each student submitted their assignment was calculated by subtracting the submission time from the assignment deadline. Because the distribution

of submission times was, as predicted, highly negatively skewed, with few students submitting early and increasing numbers of submissions as the deadline approached, a natural log transformation was performed.

The submission-time data were then group-mean centred around the mean submission time for each assignment. Group-mean centring was used rather than the more usual grand-mean centring, because we could not be certain that staff who set up the electronic assignments had entered the correct deadline. In many cases, it appeared that all students on a course had submitted a week or more before the deadline, which was implausible and does not fit with the observation that student behaviour follows a hyperbolic curve. This means that submission times, calculated by subtracting the actual time of submission from the deadline as set in the electronic system, are not reliable. Centring the submission times by assignment recalculates each student's submission time for an assignment so it is relative to submissions by other students on the same assignment. We reversed the sign so that negative numbers represented students who submitted earlier than average for that assignment and positive numbers represented those who submitted later than average – we will call this measure 'procrastination'. This removes the reliance on knowing the actual deadline, while still allowing students to be characterised as early or late submitters. The consequence of this decision is that the predictor variable 'procrastination' represents relative delay, compared with other students doing the same assignment, rather than absolute delay.

The criterion (dependent) variable was mark (numerical grade), on a percentage scale with a pass mark of 40%.

Students

Records were each associated with a student identity code. In total 21691 unique students were included in the study, with an average of 3.4 submissions each (range 1-27).

Assignments

2335 individual assignments were included in the study. Some assignments were set across several different cohorts with a similar title, but were coded as different assignments because the task and marking schemes may have differed.

Courses

There were 1272 courses (modules within a programme). Markers and types of assignments often differ within the same course.

Disciplines

Records were coded into 21 cognate discipline groupings based upon the courses owned by academic teams. The discipline groupings were: Psychology, Social Sciences, Social Work, Culture (History, American Studies), English, Languages, Geography, Tourism, Environment, Sport, Chemistry, Biology, Healthcare Sciences (Biomedical Sciences, Pharmaceutical Sciences), Pharmacy, Nursing, Education, Law, Business, Computing, Media, and Performing Arts.

Statistical analysis

Before examining the relationship between submission time and marks, it is important to recognise that the data set is not homogenous. Each data point represents one assignment submission by an individual student, but to treat these data points as

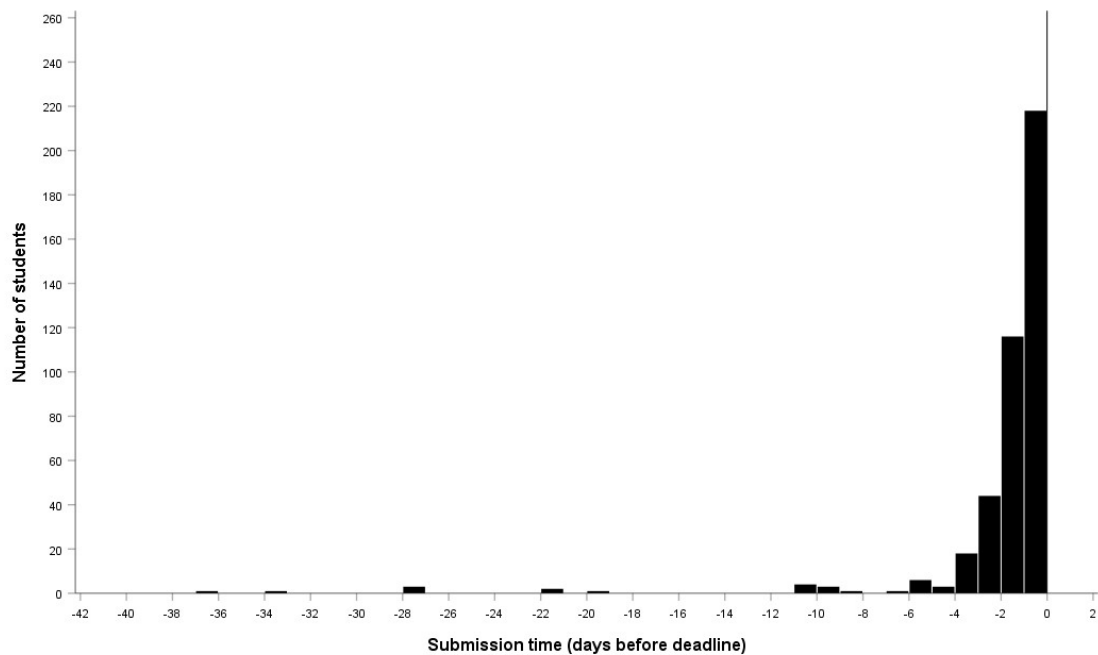
independent would be erroneous. There are several layers of structure and groupings across data points. For example each student will have contributed several data points, by submitting different assignments at different times. Data points can also be grouped by different assignments, by different courses, or even by different subject disciplines. Other groupings may exist which cannot be discerned from the existing data, such as submissions marked by a particular marker, or assignments of particular types. Because of the existence of these groupings within the data, we chose to analyse the data using multilevel modelling. This allowed some of the hierarchical structure to be taken into account (Hox 2010), to evaluate the extent to which each of the four main forms of structure: student, assignment, course, and discipline, accounted for the variance in marks, before then determining to what additional extent procrastination also explained marks.

Results

Descriptive statistics

The distribution of submission times was highly skewed. The median time of submission was 24 hours before the deadline. Most assignments showed the classic hyperbolic shape (König and Kleinman 2005; Howell et al. 2006), with increasingly more students submitting as the deadline approached, and fewer students submitting very many days in advance. An example is shown at Figure 1; the student who submitted closest to the deadline submitted less than three minutes beforehand.

Figure 1: Submission times for the largest assignment in the sample (business management)



Unconditional variance components models

All analysis was carried out in MLwiN 3.02 (Charlton et al. 2017; Rasbash et al. 2017), and R 3.6.0 (R Core Team 2019) using the lme4 package (Bates et al. 2015).

In the first stage of investigation, the students' mark was entered as the outcome variable in a two-level hierarchical model, using the iterative generalised least squares method of estimation, first with students as the level 2 random factor, then with assignments as the level 2 random factor, then with course and finally with discipline (models 1-4, respectively), to establish the amount of variance in marks accounted for by each alternative structure.

Table 1: Variance components models for two-level hierarchical structures with marks as the outcome variable

| Model | 1 | 2 | 3 | 4 |
|-----------------------|---------------------------|--------------------------|----------------------------|--------------------------|
| Random effect | <i>Student</i> | <i>Assignment</i> | <i>Course</i> | <i>Discipline</i> |
| VPC† (%) | 28.8 | 23.5 | 15.0 | 2.8 |
| Deviance | 596360 | 585981 | 593100 | 602919 |
| Likelihood ratio test | t(19610)=737.8, p<.001 | t(2063)=339.1, p<.001 | t(1220.5)=319.5, p<.001 | t(19.7)=101.1, p<.001 |

* Significance calculated using lmerTest package following the Satterthwaite method, see Luke (2017)

† Variance partition coefficient

Likelihood tests determined that all of these 2-level models made a significant contribution to explaining the variability in marks, as seen in Table 1. Clear evidence of multilevel structure can be seen by looking at the variance partition coefficient (VPC), which is here also equivalent to the intraclass correlation coefficient. When the level 2 random effect in the model was the student, the intraclass correlation between submissions was 0.288, so marks awarded to the same student across different assignments correlated. To put it another way, ~29% of the variance in marks can be attributed to individual differences between students. The remaining 71% of the variance in marks is due to other situational factors which differ from one assignment submission to the next. Some of these situational factors can be seen in the other three variance components models. With assignments as the random effect, 23.5% of the variance in marks is explained by differences between assignments. Courses and subject disciplines were responsible for smaller amounts of the variance in marks, 15% and 3%, respectively. Differences between disciplines explained relatively little of the variance in marks.

In case some of this variance is shared, further unconditional models were tested (Table 2; the first two models from Table 1 are repeated here as a baseline comparison). First we tested a cross-classified model including both students and assignments. This

model 5 accounted for significantly more variance in marks, but the proportions of variance explained by students (28%) and assignments (19%) were only slightly lower than those in the separate models, suggesting that these were largely independent sources of variance.

In model 6 we tested whether nesting assignments in courses explained more of the variance than assignments alone. In this model 14% of the variance in marks was explained by assignments and 10% by courses. Adding courses substantially reduced the amount of variance explained by assignments, suggesting that some of the apparent differences in assignments were actually due to differences in courses, where the variance barely dropped. We then added students to this model in a hierarchical cross-classified model (7). Again the proportion of variance explained by students remained the same, but the variance explained by assignments and courses dropped. Finally we added disciplines in model 8, to test whether assignments nested in courses nested in disciplines explained more of the variance in marks than model 6. In this model 13% of the variance was explained by assignments, 8% by courses and only 2% by disciplines.

Table 2: Variance components models for cross-classified and three-level hierarchical structures with marks as the outcome variable

| | Model | Deviance | Random factor | VPC (%) | Compared with | χ^2 | df |
|---|--------------------------|----------|---------------|---------|---------------|----------|----|
| 1 | Hierarchical 2 level | 596360 | Students | 29 | | | |
| 2 | Hierarchical 2 level | 585981 | Assignments | 24 | | | |
| 5 | Cross-classified 2 level | 575583 | Students | 28 | Model 1 | 20777* | 1 |
| | | | Assignments | 19 | Model 2 | 10398* | 1 |
| 6 | Hierarchical 3 level | 585744 | Assignments | 14 | Model 2 | 296* | 1 |
| | | | Courses | 10 | | | |
| 7 | Cross-classified 3 level | 505552 | Students | 28 | Model 5 | 159* | 1 |
| | | | Assignments | 12 | Model 6 | 10320* | 1 |
| | | | Courses | 7 | | | |
| 8 | Hierarchical 4 level | 585685 | Assignments | 13 | Model 6 | 59* | 1 |
| | | | Courses | 8 | | | |
| | | | Disciplines | 2 | | | |

* Significant at $p < .001$

It was not possible to compute a cross-classified 4-level model to include students as well, because it failed to converge. However, we would expect that ~28% of the variance explained would be due to students, and the proportions explained by assignments, courses and disciplines would be expected to drop by perhaps 1% each as was the case in model 7. Table 2 shows that each model explains significantly more variance than the previous ones, with students always providing the primary source of variance in marks.

Inferential statistics

Since we were unable to calculate a full model with random effects for students cross-classified with assignment nested in courses nested in disciplines, we decided to separate the analysis of student-level effects from the analysis of assignment-level effects. Model 1 is the empty model used to capture student-level effects. Since model 8 (the full hierarchy of assignments nested in courses nested in disciplines) captured significantly more of the variance in marks than model 2 or model 6, we also proceeded using model 8 as an empty model. The second stage of analysis involved adding submission time as a predictor, first only with random intercept and fixed slope, then secondly also allowing slopes to vary randomly, to determine which models best accounted for the variance in marks. This analysis was carried out in R 3.6.0, using Nakagawa et al.'s (2017) method for estimating the effect size, R^2 , for multilevel models.

Table 3 shows the addition of procrastination as an explanatory variable to model 1, first with random intercept and then allowing random slope. Both models 1a and 1b are a significant improvement on the unconditional model, with the random slope model explaining significantly more variance. This confirms that procrastination significantly explains marks, with the relationship between procrastination and marks varying depending on the student.

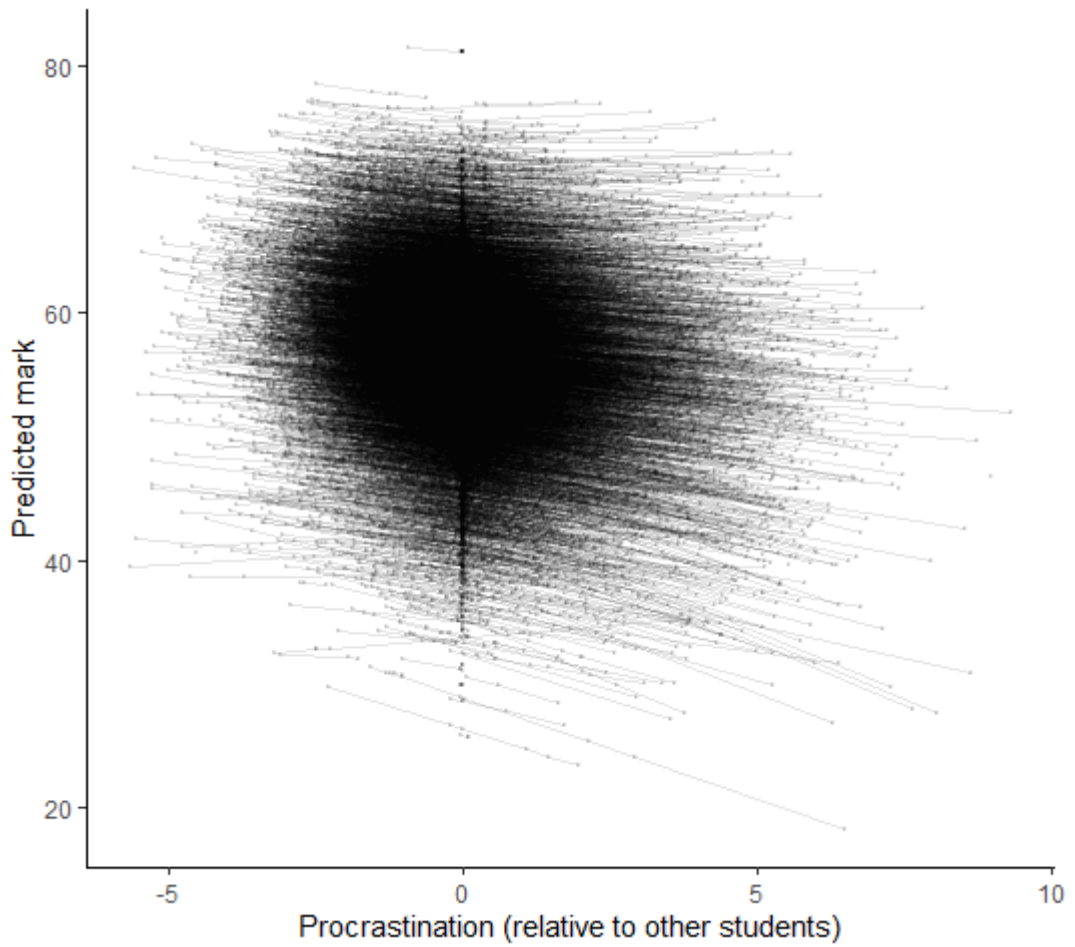
Table 3: Inferential statistics for the model fit for submission time with students as a random factor, each model compared with the previous one

| Model | | Deviance | χ^2 | df | R ² | | VPC (%) |
|-------|----------------|----------|----------|----|----------------|-------|---------|
| | | | | | marg. | cond. | |
| 1 | Student (null) | 596360 | | | - | 0.288 | 29 |
| 1a | Intercept | 596143 | 218* | 1 | 0.003 | 0.287 | |
| 1b | Slope | 596114 | 29* | 2 | 0.003 | 0.292 | |

* Significant at $p < .001$

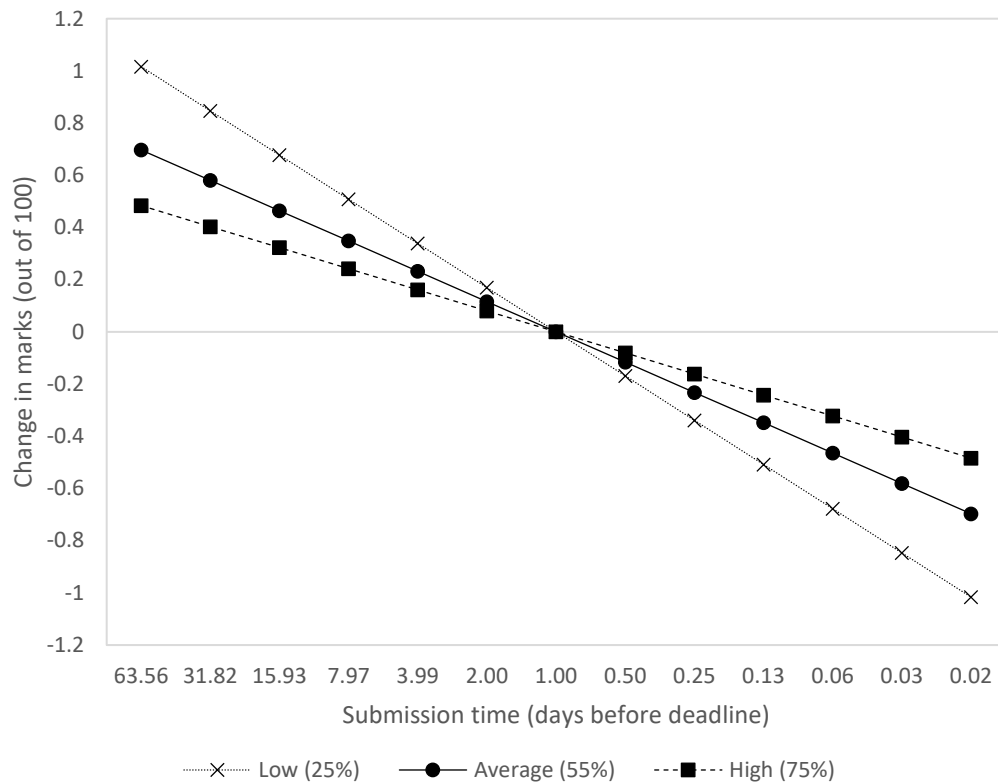
The covariance of 0.93 between slope and intercept was significant and positive indicating a fanning-out effect, so that students with lower marks had steeper (more negative) slopes, as seen in Figure 2.

Figure 2: Marks predicted from procrastination level for each student by Model 1b



To make it easier to see the overall pattern, we entered student average mark and its interaction with procrastination into the equation, to produce the schematic representation in Figure 3.

Figure 3: Change in marks depending on submission time for students with very low average marks (a fail at 25%), average marks (lower second class at 55%), and first class marks (75%), for an average assignment



Since this is looking at repeated submissions across the same student, this tells us that there is a very slight trend for greater procrastination (submission nearer the deadline) to be associated with poorer marks, and that the drop in marks is slightly higher for students who have lower marks than those with higher marks. However, the drop in marks on an average assignment between submitting one week early and half an hour before the deadline is only around 1%.

We also examined at the relationship between procrastination and marks in the context of assignments nested in courses nested in disciplines (model 8). This was a between-students comparison, since here we were looking at whether later submissions on a

particular assignment (where each submission is made by a different student) were associated with lower marks. We tested whether the relationship between submission time and mark varies by assignment, by course, and by subject discipline. Table 4 shows that model 8b, with random slopes permitted for all the levels (assignment, course and discipline) accounted for significantly more of the variance in marks. This means that there was a significant relationship between submission time and marks, and that this relationship varied depending on assignment, course and discipline, however this relationship was again very weak, accounting for less than 1% of the variance in marks.

Some assignments had a shallower slope than others, suggesting that procrastination was less important in some assignments than others. The relationships did not vary in any systematic way according to marks: the covariances between intercept and slope were not significant. The pattern for courses and for disciplines was similar to that for assignments. For academic disciplines, those with the strongest relationship between procrastination and marks were Environment and Geography, and those with the least were Business, Law, and Nursing. There was little evidence of any cognate subject clustering.

Table 4: Inferential statistics for the model fit for submission time with assignments nested in courses nested in disciplines as random factors, each model compared with the previous one

| Model | | Deviance | χ^2 | df | R ² | | VPC (%) |
|-------|--|----------|----------|----|----------------|-------|---------|
| | | | | | marg. | cond. | |
| 8 | Assignment, Course & Discipline (null) | 585685 | | | - | 0.236 | 23.6 |
| 8a | Intercept | 585029 | 655* | 1 | 0.003 | 0.243 | |
| 8b | Slope (all) | 584948 | 82* | 5 | 0.003 | 0.248 | |

* Significant at $p < .001$

Discussion

We expected, based on previous literature (Steel, Brothan, and Wambach 2001; Moon and Illingworth 2005; Rotenstein, Davis, and Tatum 2009; Levy and Ramim 2012; Hensley 2014; Arnott and Dacko 2014; Kim and Seo 2015; Cerezo et al. 2016), that assignment submission time (as a proxy measure of procrastination) would predict marks, such that later submissions would be associated with lower marks. This prediction was confirmed, with submission time significantly predicting marks, however effect sizes were very small. This was the case both within-students (longitudinal) and between-students (cross-sectional). In all the models the amount of variance in marks which was explained by procrastination was negligible, at less than 1%, and as low as 0.3% when student-level factors were taken into account.

Inspection of the empty models first established that multilevel modelling was necessary, since marks varied significantly according to student, assignment, course, and discipline, so each submission could not be treated independently of others. The main source of variance in marks (at 28-29%) was students, showing a certain consistency in the marks that an individual student received across assignments. It was reassuring that this was the main source of variance, and was much higher than variance in marks due to assignments, courses or disciplines, since variation in marks between assignments can potentially aggregate into differences large enough to affect students' degree classifications, creating unfairness in the system. There is no evidence of such an issue in this data set: we found that as the context became more general (from assignments to courses to disciplines), variance in marks reduced. The variance in marks from one discipline to another was relatively small at 2%, and could have been

lower if we could also have excluded student-level factors (that some disciplines may have a greater number of high-performing students, for example due to higher entry qualifications). These findings suggest that a student's final degree classification is unlikely to be heavily influenced by their choice of discipline or course options, even if some assignments tend to be 'harder' or 'easier'.

The within-students prediction was tested in a multilevel model with students as a random-effects factor. This showed a significant but very weak negative relationship between submission time and mark. On average, when a student submitted an assignment relatively late, that student obtained marginally lower marks than on assignments when they submitted earlier. The best fit model was a random-slopes model, where the strength of the relationship between submission time and marks varied by student. This could be because submission time is less of a measure of procrastination for some students than others, or it could be because procrastination is less detrimental for some students than others, as claimed by Chu and Choi (2005). However, there was also a systematic pattern of variation between students. The negative slope was a little steeper for students with lower average marks than for students with higher average marks, so later submission was better at predicting a drop in marks for lower-scoring students than for higher-scoring students.

A direct causal relationship between submission time and marks can be dismissed as implausible, since it is unlikely that the student would have obtained higher marks on a particular assignment if they had stopped work and handed in that assignment earlier. It is more likely that on the occasions that students submitted relatively late, they did so because they had insufficient time to complete the

assignment to their usual standard, and were still trying to improve it towards the end. This could be due to procrastination, or other pressures beyond the student's control such as work or technical issues. Since whatever caused the late submission was more detrimental to students scoring at the lower end of the marking scale, it is possible that aspects of an assignment which contribute most to marks were completed earlier by first class-scoring students, so that running out of time was less costly than for students scoring at the low end. These speculations would need to be investigated in further research, however the effect sizes involved were very small, with substantially later relative submission necessary in order to observe as much as a 1% drop in marks.

The between-students prediction was tested in a multilevel model with assignments nested in courses nested in disciplines as random factors. Here we found that later submissions in all these contexts attracted significantly lower marks, though the effect size was again very small. The fully random-slope model was the best model, so that the relationship between submission time and marks varied by assignment, by course, and by discipline. For example, in Business the negative slope was shallow, whereas in Environment and Geography, the slope was steeper, so students who submitted close to the deadline in Environment or Geography were more likely to risk lower marks than those submitting late in Business. There was no evidence of delay being less detrimental in 'easier' assignments, courses or disciplines (those where more students got high marks): just a slight tendency for delay to be detrimental across all assignments on average.

Although this large scale study has confirmed the negative relationship between submission time and grades found by Steel, Brothan, and Wambach (2001), Moon and

Illingworth (2005), Rotenstein, Davis, and Tatum (2009), Levy and Ramim (2012), Hensley (2014), Arnott and Dacko (2014), and Cerezo et al. (2016), the small effect size and unreliability of the association for any particular student or assignment suggests that submission time would not be a useful metric for identifying at-risk students.

Kim and Seo's (2015) meta-analysis of the existing studies at that time suggested that behavioural measures were more promising than self-reported measures of procrastination for showing a relationship with performance, however this was based on a small number of studies which, compared with our findings, seem anomalous. The medium to large relationships found by Steel et al. (2001) and Howell et al. (2006) would be very unusual. Even in the studies we reviewed in the introduction to this paper, small correlations (0.1 to 0.2) were more common, with larger samples rendering these significant. Our large-scale testing suggests the average correlation is much smaller (around 0.08 at most) which does suggest a possible file drawer effect (Rosenthal 1979). Since most marking interfaces allow markers to view both submission time and marks, it is possible that many markers have observed or even statistically tested possible associations between them, with those obtaining results that showed no significant relationship proceeding no further, so only the most striking results have made it into the public domain.

The main limitation of our study is that it looked at only one university, where nearly all the courses were on-campus and assessed by high-stakes coursework assignments. There may be other contexts for which the relationship may be stronger and more consistent. However, what is definitely clear from our study is that, in most contexts, it is unlikely that the relationship between submission time and grade will be

useful for predictive purposes, and there is no reason to believe that findings from previous literature will replicate in different assignments, courses, disciplines, or universities. It would be useful to carry out further work to narrow down in what conditions the relationship may be stronger, for example identifying if there are any common factors shared by assignments for which procrastination is more strongly associated with low marks.

A second limitation of our study is that we used submission time as a proxy measure of procrastination; the small effect sizes found may be because submission time is not a very good measure of procrastination, or may be because procrastination is not very strongly associated with performance. Further research is needed to validate submission time as a measure of procrastination.

In conclusion, in this large scale study we have demonstrated that, contrary to more promising small-scale findings, submission time alone is not a useful predictor of mark for a student on an assignment, as effect sizes are extremely small and the relationship between submission time and marks varies depending on both student and assignment context.

References

Agudo-Peregrina, A. F., S. Iglesias-Pradas, M. A. Conde-González, and A. Hernández-García. 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, 542-550. doi: 10.1016/j.chb.2013.05.031

- Arnott, D., and S. Dacko. 2014, June 4. “*Time of Submission: An Indicator of Procrastination and a Correlate of Performance on Undergraduate Marketing Assignments.*” Paper presented at the 43rd European Marketing Academy Conference: Paradigm Shifts & Interactions, Valencia, Spain.
- Baker, R. S., D. Lindrum, M. J. Lindrum, and D. Perkowski. 2015, June 26-29. “*Analyzing Early At-Risk Factors in Higher Education E-Learning Courses.*” Paper presented at the 8th International Conference on Educational Data Mining, Madrid, Spain.
- Bates, D., M. Maechler, B. Bolker, and S. Walker. 2015. “Fitting Linear Mixed-Effects Models Using lme4.” *Journal of Statistical Software* 67 (1): 1-48. doi: 10.18637/jss.v067.i01
- Cerezo, R., M. S. Sanchez-Santillana, M. Paule-Ruiz, and J. C. Núñez. 2016. “Students' LMS Interaction Patterns and Their Relationship with Achievement: A Case Study in Higher Education.” *Computers and Education* 96, 42-54. doi: 10.1016/j.compedu.2016.02.006
- Charlton, C., J. Rasbash, W. J. Browne, M. Healy, and B. Cameron. 2017. “*MLwiN Version 3.00.*” Centre for Multilevel Modelling, University of Bristol, UK.
- Choi, J. N., and S. V. Moran. 2009. “Why not procrastinate? Development and Validation of a New Active Procrastination Scale.” *The Journal of Social Psychology* 149 (2): 195–212. doi: 10.3200/socp.149.2.195-212
- Chu, C. A. H. and J. N. Choi. 2005. “Rethinking Procrastination: Positive Effects of ‘Active’ Procrastination Behavior on Attitudes and Performance.” *The Journal of Social Psychology* 145 (3): 245–264. doi: 10.3200/socp.145.3.245-264.
- Colvin, C., T. Rogers, A. Wade, S. Dawson, D. Gasevic, S. Buckingham Shum, K. Nelson, S. Alexander, L. Lockyer, G. Kennedy, L. Corrin, and J. Fisher. 2015. “*Student Retention and Learning Analytics: A Snapshot of Australian Practices*

and a Framework for Advancement.” Australian Government Office for Learning and Teaching.

- Corkin, D. M., S. L. Yu, and S. F. Lindt. 2011. “Comparing Active Delay and Procrastination from a Self-Regulated Learning Perspective.” *Learning and Individual Differences* 21, 602–606. doi: 10.1016/j.lindif.2011.07.005
- Dawson, S, D. Gašević, G. Siemens, and S. Joksimovic. 2014. “Current State and Future Trends: a Citation Network Analysis of the Learning Analytics Field.” *Proceedings of the Fourth International Conference on Learning Analytics And Knowledge (LAK '14)*. ACM, New York, NY, USA, 231-240. doi: 10.1145/2567574.2567585
- De Freitas, S., D. Gibson, C. du Plessis, P. Halloran, E. Williams, M. Ambrose, I. Dunwell, and S. Arnab. 2014. “Foundations of Dynamic Learning Analytics: Using University Student Data to Increase Retention.” *British Journal of Educational Technology* 46 (6): 1175–1188. doi: 10.1111/bjet.12212
- Dewitte, S. and H. C. Schoewenburg. 2002. “Procrastination, Temptations, and Incentives: the Struggle between the Present and the Future in Procrastinators and the Punctual.” *European Journal of Personality* 16, 469-489. doi: 10.1002/per.461
- Gafni, R. and N. Geri. 2010. “Time Management: Procrastination Tendency in Individual and Collaborative Tasks.” *Interdisciplinary Journal of Information, Knowledge, and Management* 5, 115–125. doi: 10.28945/1127
- Gašević, D., S. Dawson, T. Rogers, and D. Gašević. 2016. “Learning Analytics Should Not Promote One Size Fits All: The Effects of Instructional Conditions in Predicting Academic Success.” *Internet and Higher Education* 28, 68-84. doi: 10.1016/j.iheduc.2015.10.002

- Hensley, L. C. 2014. "Reconsidering Active Procrastination: Relations to Motivation and Achievement in College Anatomy." *Learning and Individual Differences* 36, 157-164. doi: 10.1016/j.lindif.2014.10.012
- Howell, A. J., D. C. Watson, R. A. Powell, and K. Buro. 2006. "Academic Procrastination: The Pattern and Correlates of Behavioural Postponement." *Personality and Individual Differences* 40 (8): 1519–1530. doi: 10.1016/j.paid.2005.11.023
- Hox, J. J. 2010. *Multilevel Analysis: Techniques and Applications*. 2nd ed. New York, US: Routledge/Taylor & Francis
- Jo, I.-H., Y. Park, M. Yoon, and H. Sung, H. 2016. "Evaluation of Online Log Variables that Estimate Learners' Time Management in a Korean Online Learning Context." *The International Review of Research in Open and Distributed Learning* 17(1). doi: 10.19173/irrodl.v17i1.2176
- Kim K and E. Seo. 2015. "The Relationship between Procrastination and Academic Performance: A Meta-Analysis." *Personality and Individual Differences* 82, 26-33. doi: 10.1016/j.paid.2015.02.038
- Kim, S., S. Fernandez, and L. Terrier. 2017. "Procrastination, Personality Traits, and Academic Performance: When Active and Passive Procrastination Tell a Different Story." *Personality and Individual Differences* 108, 154–157. doi: 10.1016/j.paid.2016.12.021
- Klingsieck, K. B. 2013. "Procrastination: When Good Things Don't Come to Those who Wait." *European Psychologist* 18, 24-34. doi: 10.1027/1016-9040/a000138
- König, C. J. and M. Kleinman. 2005. "Deadline Rush: A Time Management Phenomenon and its Mathematical Description." *The Journal of Psychology* 139 (1): 33-45. doi: 10.3200/jrlp.139.1.33-45

- Levy, Y. and M. M. Ramim. 2012. "A Study of Online Exams Procrastination Using Data Analytics Techniques." *Interdisciplinary Journal of E-Learning and Learning Objects* 8, 97–113. doi: 10.28945/1730
- Luke, S. G. 2017. "Evaluating significance in linear mixed-effects models in R." *Behavior Research Methods* 49, 1494-1502. doi: 10.3758/s13428-016-0809-y
- McElroy, B. W. and B. H. Lubich. 2013. "Predictors of Course Outcomes: Early Indicators of Delay in Online Classroom." *Distance Education* 34 (1): 84–96. doi: 10.1080/01587919.2013.770433
- Moon, S. M. and A. J. Illingworth. 2005. "Exploring the Dynamic Nature of Procrastination: A Latent Growth Curve Analysis of Academic Procrastination." *Personality and Individual Differences* 38 (2): 297–309. doi: 10.1016/j.paid.2004.04.009
- Nakagawa, S., P. Johnson, P., and H. Schielzeth. 2017. "The Coefficient of Determination R^2 and Intra-Class Correlation Coefficient from Generalized Linear Mixed-Effects Models Revisited and Expanded." *Journal of the Royal Society Interface* 14 (134). doi: 10.1098/rsif.2017.0213
- Oliphant, T. E. 2006. *A Guide to NumPy*. USA: Trelgol Publishing.
- R Core Team. 2019. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>
- Rasbash, J., F. Steele, W. J. Browne, and H. Goldstein. 2017. *A User's Guide to MLwiN, v3.01*. Centre for Multilevel Modelling, University of Bristol.
- Rosenthal, R. 1979. "The File Drawer Problem and Tolerance for Null Results." *Psychological Bulletin* 86 (3): 638-641. doi: 10.1037/0033-2909.86.3.638

- Rotenstein, A., H. Z. Davis, and L. Tatum. 2009. "Early Birds versus Just-in-Timers: The Effect of Procrastination on Academic Performance of Accounting Students." *Journal of Accounting Education* 27, 223–232. doi: 10.1016/j.jaccedu.2010.08.001
- Rozenal, A., E. Forsell, A. Svensson, D. Forsström, G. Andersson, and P. Carlbring. 2015. "Differentiating Procrastinators from Each Other: A Cluster Analysis." *Cognitive Behaviour Therapy* 44 (6): 480-490. doi: 10.1080/16506073.2015.1059353
- Siemens (2010) Definition provided on 22nd July 2010 for the 1st International Conference on Learning Analytics and Knowledge, Banff, Alberta, February 27–March 1, 2011 Accessed from: <https://tekri.athabascau.ca/analytics/about>
- Solomon, L.J. and E. D. Rothblum. 1984. "Academic Procrastination: Frequency and Cognitive-Behavioral Correlates." *Journal of Counseling Psychology* 31 (4): 503–509. doi: 10.1037/0022-0167.31.4.503
- Steel, P. 2007. "The Nature of Procrastination: A Meta-Analytic and Theoretical Review of Quintessential Self-Regulatory Failure." *Psychological Bulletin* 133 (1): 65–94. doi: 10.1037/0033-2909.133.1.65
- Steel, P., T. Brothen, and C. Wambach. 2001. "Procrastination and Personality, Performance, and Mood." *Personality and Individual Differences* 30 (1): 95-106. doi: 10.1016/S0191-8869(00)00013-1
- Thille, C., E. Schneider, R. F. Kizilcec, C. Piech, S. A. Halawa, and D. K. Greene. 2014. "The Future of Data-Enriched Assessment." *Research & Practice in Assessment* 9 (2): 5–12.
- Tice, D. M., and R. F. Baumeister. 1997. "Longitudinal Study of Procrastination, Performance, Stress, and Health: The Costs and Benefits of Dawdling." *Psychological Science* 8, 454-458. doi: 10.1111/j.1467-9280.1997.tb00460.x