

Impact of Student Choice of Content Adoption Delay on Course Outcomes

Lalitha Agnihotri
McGraw-Hill Education
2 Penn Plaza
New York, NY, USA
lalitha.agnihotri
@mheducation.com

Alfred Essa
McGraw-Hill Education
281 Summer Street, 7th Floor
Boston, MA, USA
alfred.essa
@mheducation.com

Ryan Baker
University of Pennsylvania
3700 Walnut St.
Philadelphia, PA, USA
rybaker@upenn.edu

ABSTRACT

It is difficult for a student to succeed in a course without access to course materials and assignments; and yet, some students delay up to a month in obtaining access to these essential materials. Students delay buying material required for their course due to multiple reasons. Out of a concern for students with limited financial resources, some publishers offer a period of free courtesy access. But this may lead to students having access later in the course but then having a lapsed period until they pay for the materials after the courtesy access period ends. Not having key course materials early on probably hurts learning, but how much? In this paper, we investigate the question, "Does lack of access to instructional material impact student performance in blended learning courses?" Specifically, we analyze students who purchased and obtained access to online content at different points in the course. We determine that both types of failure to obtain access to course materials (delaying in signing up for the product, or signing up for a free trial and letting the trial period lapse without purchasing the materials) are associated with substantially worse student outcomes. Students who purchased the product within the first few days of class had the best scores (median 77). Those who waited two weeks before accessing the product did the worst (median 56, effect size Cliff's Delta=0.31¹). We conclude with a discussion of possible interventions and actions that can be taken to ameliorate the situation.

CCS Concepts

•Information systems → Data mining; •Applied computing → Education; E-learning;

¹For Cliff's Delta a small effect size is around 0.147, a medium effect size around 0.33, and a large effect size around 0.474.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

LAK 2017 Vancouver, BC, Canada

© 2017 ACM. ISBN 123-4567-24-567/08/06...\$15.00

DOI: 10.475/123_4

Keywords

effect size, procrastination, content adoption delay, performance

1. INTRODUCTION

There is evidence that students benefit when they start their work early rather than waiting to start or procrastinating [4, 6]. However, despite the evidence for benefits from getting a prompt start to a course, there is emerging evidence that many learners do not even purchase their course materials until one or even two weeks after the course has started. In this paper, we study the impact of this student choice on course outcomes, and propose interventions that may have the potential to reduce negative outcomes stemming from this choice.

More specifically, in this paper, we investigate the question 'Does lack of access to textbooks and digital instruction resource significantly affect learning performance?' Students delay getting access to material required for their course for all sorts of reasons. Not having key course materials early on probably hurts learning, but how much? There are multiple reasons to look into this question: Many instructors believe that it has an effect when a student delays getting the materials required for their course. But how much do their grades suffer? And how long can a student delay on this before there is a detrimental effect? And if it does impact outcomes, what interventions can we apply in order to insure that student indeed do get access to material in a timely fashion? We will discuss some ways that it may be possible to intervene and address the issue.

2. LITERATURE REVIEW

There are multiple reasons why students may delay in purchasing their course materials. For some students it may simply be procrastination [10]. 87% of the 13,000 high school and college students surveyed by StudyMode.com admitted to procrastinating. 45% of students surveyed reported that they believe that their procrastination negatively impacts their grades on at least a fairly regular basis. Other students may be trying to decide what course to take. For some students, it simply comes down to the fact that they do not receive their financial aid check until two weeks in the semester, and they can not afford to buy materials before then. Research has shown that even opening the text book prior to the start of course is predictive of success in the course [4].

Agnihotri and Ott also determined that another form of procrastination, late registration, is associated with lower fall-to-fall student retention [2]. In addition, Levy and Rahm [8] found that students who procrastinated performed significantly worse than those who completed their work in a more timely fashion. Results of this study indicate that when it comes to online exams, over half (58%) of the students tend to procrastinate, while the rest (42%) started the exam well before the deadline and avoided procrastination.

Jayaprakash et al., determined that course success can be predicted from the student’s interaction with the learning management system [7]. Predictive models have also been developed by Civitas and deployed at a range of institutions [9]. Their predictive models were able to identify with 83% accuracy on the first day of a course the students who would successfully complete a course based on ACT scores, SAT scores and economic factors.

3. METHOD

In order to estimate the impact of student choice about when to obtain access to content, and when students purchase access to content, we look at the differences between different groups of students. Specifically, we differentiated groups of students from each other in terms of how long they chose to go without access to course materials.

We study this in the context of the Connect system [1]. Connect is an open learning environment for students and instructors in the higher education market. In this analysis, we examine this utilized about 2.6 million students who used Connect in 2015. These students were in 145,115 course sections taught by 14,000 instructors, who created 89 million assignments using about 2000 textbooks/course material packets. The majority (75%) of the students who obtain access to Connect purchase access outright. However, there is an option for students to try it for free for two weeks (termed Courtesy Access) and then convert it to full access at a later date. Of the students who opt for Courtesy Access, 80% convert to full access.

For all the students we obtained data about when they got access to Connect. Additionally we obtained the start date of the class. We use this information to compute two variables: Start delay is defined as how many days after the start of the course the student first obtained access to the online content, whether by purchasing the course or obtaining courtesy access. Conversion delay, defined only for those students who obtained courtesy access and then eventually purchased access, is the number of days between when their courtesy access period started and when they converted to full access. Since the courtesy access period is two weeks, students with a conversion delay of two weeks or less have a conversion delay of zero. We also obtained data on students’ assignment scores and quiz scores, and computed their final scores for the class based on this data. We compute these scores in two ways. The first, termed "ScoreCompleted", is a strict average of all the scores students have received on the assignments/quizzes etc. that they submitted. The second one, termed "ScoreAll", shows the score with the impact of missed assignments factored in. In other words, if a student failed to do two assignments due to not having the materials for two weeks, "ScoreAll" will directly penalize them but "ScoreCompleted" will not.

A quick analysis of our data showed that these variables were largely non-normal. As such, we compared the scores

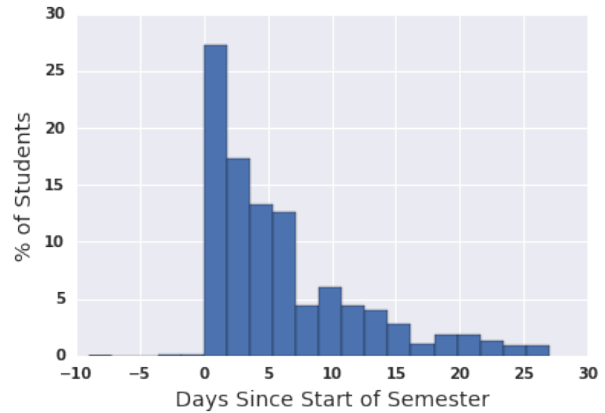


Figure 1: Histogram of acquiring access relative to start of semester.

of students who obtain access of the book at different points using the Cliff’s Delta effect size measure [5]. The Cliff’s Delta statistic is a non-parametric effect size measure that quantifies the amount of difference between two groups of observations. This effect size measure is used for non-normal distributions; an analogue for normal distributions is Cohen’s D. Cliff’s Delta was chosen for its particularly high robustness to unusual data distributions; other alternatives such as Algina’s D control for outliers but not for bimodality or extremely high skew.

4. CONVERSION DATES AND PERFORMANCE

Our data set consisted of:

- 2.6 million students in 145,115 sections in 2015, who made a total 3.2 million purchases
- 2.4 million (75%) outright purchases (i.e. without first signing up for a Courtesy Access period)
- 818k (25%) Courtesy Access trials: 633k (77% of 818k) purchases after trial, 185k (23% of 818k) trials without purchases

Figures 1 and 2 show the histogram of getting access to the course material relative to the start of the semester (start delay on the X axis vs. counts on the Y axis) and the conversions relative to the start of courtesy access (conversion delay on the X axis vs. counts on the Y axis). We track up to 30 days after the start of courtesy access in our data. 47% of the student get access to content (full or courtesy access) in the first 4 days of the semester. Another 38% happens between the 5th-12th days and finally 14% occur 12 or more days after the semester starts. Very few students obtain access to course materials prior to the official start date of the course, a contrast to the results presented in [4]. This is largely because the way the courses are set up; students typically receive the link to obtain course materials on the first day of class from the instructor.

In terms of conversions from Courtesy Access to full (paid) access, 54.0% of conversions happen in less than 14 days, a time window where the student has no lapse in their access to content. Another 20.0% conversions happen in 14-16 days, suggesting fairly limited time lapse and fairly limited disruption to the student’s studies. In fact, 14 days is the

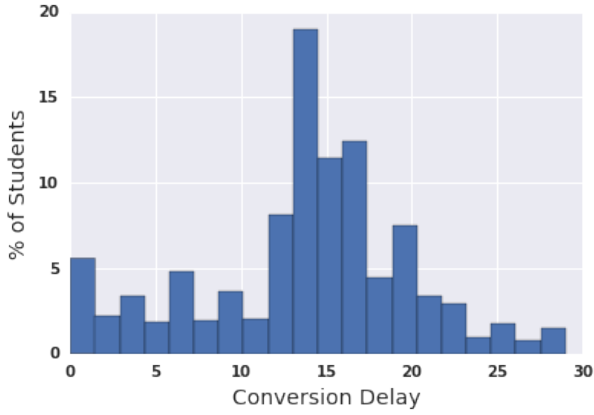


Figure 2: Histogram of conversion dates relative to start of Courtesy Access.

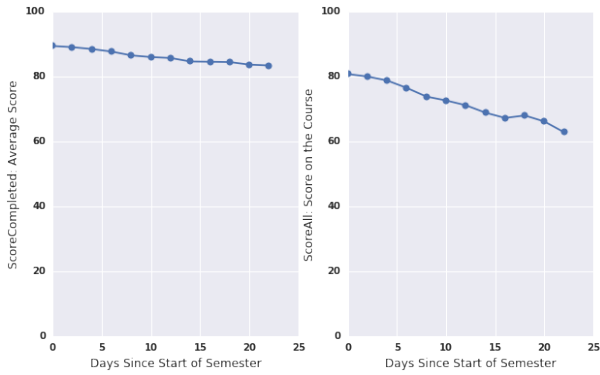


Figure 3: ScoreCompleted and ScoreAll relative to start of semester.

modal day for conversion. However, a sizable 26.0% of conversions happen after 16 days of the start of Courtesy Access. And 7% of conversions occur more than 21 days after the start of Courtesy Access, indicating that the student is without access for the whole week (see Figure 2).

Figures 3 and 4 show the student assignment/quiz scores relative to the two delays we have talked about. In each figure, the first plot shows the ScoreCompleted vs. the time delay and the second one shows the ScoreAll which considers the missed assignments as well. The graphs in figure 3 show that performance on the completed assignments (ScoreCompleted) drops a bit for students who delay in getting access once the semester starts. Figure 4 shows the same but with respect to getting full access (conversion delay) to the product after starting free courtesy access. A student who obtains access on the first day of the course and immediately purchases access will have a median ScoreCompleted of 89%. By contrast, a student who waits 14 days to obtain access will have a median ScoreCompleted of 84%. The ScoreAll for students who get access on the first day of the class is 81%. A student who waits 14 days to obtain access will have an average ScoreAll of 67.5%, and a student who waits a full week or more to convert to full access after their 2-week Courtesy Access period ends (i.e. 21 days after obtaining Courtesy Access) will have an average ScoreAll of 64%.

These results suggest that students who choose (for whatever reason) to not have access to course materials for a

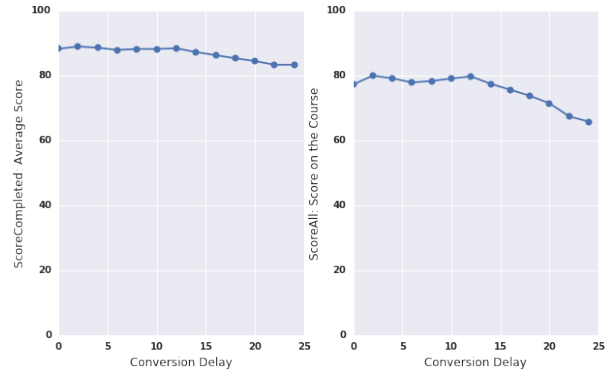


Figure 4: ScoreCompleted and ScoreAll relative to the start of the Courtesy Access.

period of time have worse outcomes, but that much of this difference (81% to 67%) is due to missing assignments rather than worse performance on the assignments they complete. This is reassuring, because it suggests that encouraging students to purchase or obtain their materials in a timely fashion has the potential to ameliorate the missed assignments problem, providing students with a chance to perform better in the course (and learn all the material). Of course, encouraging students to purchase or obtain their materials in a timely fashion will not benefit all students; for example, students who cease participation in the course for a week due to a personal or family emergency are unlikely to be benefitted. But positive impact may be possible for the students who fail to purchase or obtain their materials due to simple procrastination [6, 8]. After all, no matter how bright a student is, he or she cannot successfully complete an assignment that he or she does not have access to.

Of course, many of the students who have a start delay will also have a purchase delay. The same factors that lead to one may lead to the other. The relationship between start delay and conversion delay, and the associated scores, are shown in figure 5. The x-axis shows the start delay for getting access to the online content. The y-axis shows the conversion delay, the time the student delayed between obtaining Courtesy Access and purchasing full access. The size of the circle indicates the number of people in that group. The color indicates the median score of the students in that cohort. As can be seen, students have relatively better scores when the start delay is less than 4 days and the conversion delay is less than 14 days. When the start delay or conversion delay go above these numbers, the student is likely to obtain a lower score.

5. MAGNITUDE OF DIFFERENCES

In order to quantify the effects of start delay and conversion delay after signing up for courtesy access, we computed Cliff's Delta effect sizes on ScoreAll between groups of students who delayed for different amounts of time. Cliff's delta or d [5] is a measure of how often one the values in one distribution are larger than the values in a second distribution. Crucially, it is non-parametric and does not require any assumptions about the shape or spread of the two distributions. The sample estimate d is given by:

$$d = \frac{\#(x_i > x_j) - \#(x_i < x_j)}{mn} \quad (1)$$

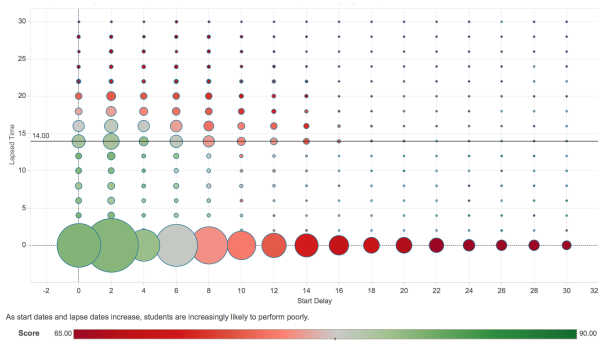


Figure 5: Heat map of scores for different start and purchase delays.

where the two distributions are of size n and m with items x_i and x_j , respectively, and $\#$ is defined as the number of times. d is linearly related to the Mann-Whitney U statistic, however it captures the direction of the difference in its sign which is important to us in this study. Cliff's delta ranges from +1 when all the values in one group are higher than the values of the other group, in the expected direction and -1 when the reverse is true. Two completely overlapping distributions will have a Cliff's delta of 0. Cliff's delta evaluates the degree of overlapping between two vectors of observations. A less raw interpretation, is to use conventional descriptors like Cohen's d (small, medium, large), which are explicitly conventional according to Cohen. For Cliff's Delta absolute value you have a small effect size around 0.15, a medium effect size around 0.33, and a large effect size around 0.50.

We computed the Cliff's delta for each of the combinations of the start delay and conversion delay. More specifically, for the start delay, we computed Cliff's delta measure for all the students scores with start delay less than or equal to a vs. start delay greater than a , where a takes values from 2 to 25. So in the above equation, we set x_i to be student scores whose start delay is less than or equal to a and x_j is the student scores for the rest of the students. We then found the start delay, a that resulted in the maximum Cliff's delta. Also, we computed the effect size for students with start delay less than or equal to a vs. students with start delay greater than b , where b takes all possible values from a to 25. So in the above equation, we set x_i to be student scores whose start delay is less than or equal to a and x_j is the student scores for students with start delay greater than b . We repeated the procedure for delay in converting to full access after obtaining 2 week courtesy access, as well. We want to find automatic cutoff points where there was maximum impact on the students' scores. We finally repeated the procedure with different combinations of start and conversion delays. To get the results we ran about 4000 different combinations of different start and conversion times to get all the different Cliff's delta.

- For students with start delay less than or equal to 12 days, the median score is 74.4% vs. students with a start delay of more than 12 days, the median score was 62.7%. Cliff's delta was 0.17.
- For students with start delay less than 3 days, the median score is 76.7% vs. students with a start delay of more than 12 days, the median score was 62.7%. Cliff's delta was 0.20.

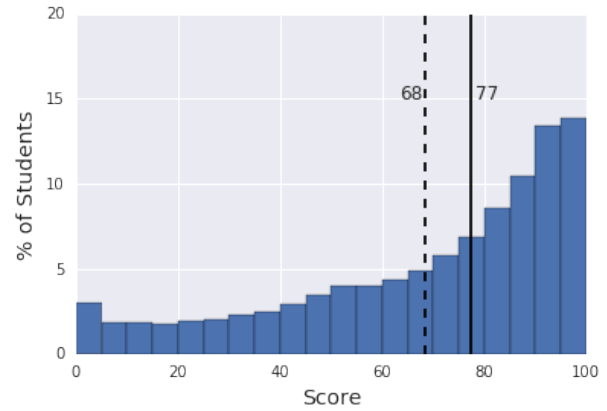


Figure 6: Score distributions of students with start delay less than 3 days and conversion delay less than 15 days.

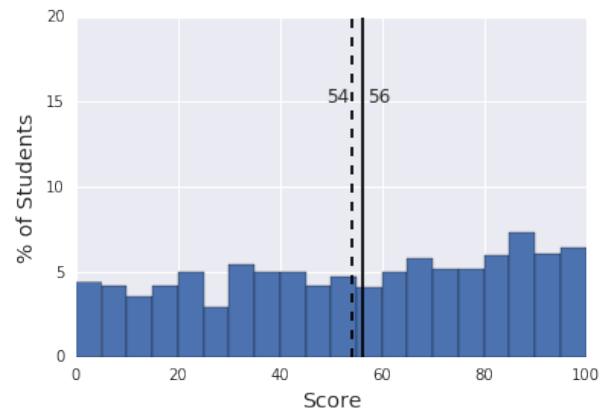


Figure 7: Score distributions of students with start delay greater than 15 days and conversion delay greater than 23 days.

- For students with conversion delay less than 19 days, the median score is 73.5% vs. students with a conversion delay of 19 days or more, the median score was 63.9%. Cliff's delta was 0.14.
- For students with conversion delay less than 16, the median score is 73.6% vs. conversion delay greater than 22, the median score was 60.4%. Cliff's delta was 0.19.

We then found automatic cut-offs for combinations of both the start delay and conversion delay:

- The Cliff's delta students with start delay less than 3 days and conversion delay less than 23 days (Median score 76.9%) vs. all other students (Median score 60.3%) is 0.25
- For varying start and conversion delays, students with start delay less than 3 days and conversion delay less than 15 days do much better (Median score 77.3%) than students who get access 15 days of the start of the semester and have a conversion time greater than 23 days (Median score 56.4%). Cliff's delta is 0.31.

Overall, then, the students who have the highest performance in their courses access the course materials within

the first few days after the start of the class. If they opt for the free courtesy access, then they are more successful if they convert to full access before they lose access to content. The worst choice is to wait for two weeks or more to obtain access to content and then let the courtesy access lapse for a week or more before converting to full access. Figures 6 and 7 show the distribution of scores for these two extremes. The odds ratio of the second group getting a score less than 60 is 2.44 and the risk ratio of getting this score and possibly failing the course is 1.68.

6. INTERVENTION POSSIBILITIES

While our results are correlational, they nonetheless show large differences in student outcomes based on when students access course materials. These findings therefore warrant intervention studies that can both validate whether these findings are causal, while testing interventions that may be able to improve student outcomes. The findings presented here suggest that there is the opportunity for improving student outcomes if we can convince students to access course materials from the beginning, and to avoid lapses in access.

One clear intervention is to simply give free access to every student. Unfortunately, as the Connect product team and project researchers need to earn money in order to eat, this solution is probably infeasible. However, to the extent that some failure to purchase course materials is due to student economic situations, such as delays in students receiving financial aid (students also need to eat), it may be possible for universities to arrange support for their students so that they can purchase materials on time. The two-week Courtesy Access period was originally designed with this in mind, but does not seem to be sufficient.

A related intervention, sometimes termed "inclusive access", is to set up a university-wide program to automatically provide all students with courtesy access to the online content at the beginning of class. If they drop the course, the content is not charged. This will help students who tend to procrastinate get access to content and facilitates coordination at the university level between when the student receives financial aid and when they are charged for the course materials.

Where this type of program is infeasible, other solutions may help students who delay in obtaining or purchasing access due to reasons such as procrastination. One approach is to work with instructors to emphasize to students the importance of getting access to the course material from the beginning. For example, it may be possible to create infographics that can be shared with instructors showing them the impact of delays in students obtaining access to content. Another potentially useful approach may be to nudge students to buy the product when the courtesy access lapses. Previous work has shown the benefits obtained from instructors sending email messages to students at risk of poorer performance, explaining why they are at risk [3].

7. DISCUSSIONS AND CONCLUSIONS

These findings indicate that it is important for students to get going quickly and avoid delay. Getting off to a fast start seems to be important for student success. One limitation to our findings, however, is that they are correlational

rather than causal. Investigating the degree to which these findings are causal, through an experimental study, will be an important step for future work.

What can we do to improve outcomes? It may be valuable to set up inclusive access, where students have free trials that last until they can be expected to receive financial aid checks. Additionally, instructors should emphasize to students that it is important to sign up for access to the course material from the beginning. Finally, students should be nudged to buy the product when the trial period lapses, in order to avoid having a period of time where they don't have access to their learning materials.

Ultimately, taking a college course without access to the learning materials is not a recipe for success. Determining which interventions can feasibly increase student access to course materials may be a valuable step towards improving student outcomes.

8. ACKNOWLEDGMENTS

The authors would like to thank Stefan Slater for Python code for computing Cliff's Delta in linear time. We would also like to acknowledge Shirin Mojarad and Nick Lewkow for their suggestions for performing analysis.

9. REFERENCES

- [1] Connect, open learning platform. <http://www.mheducation.com/highered/platforms/connect.html>. Accessed: 2016-10-09.
- [2] L. Agnihotri and A. Ott. Who are your at-risk students? using data mining to target intervention efforts. In *National Symposium on Student Retention*. The Consortium for Student Retention Data Exchanges, November 2013.
- [3] K. E. Arnold and M. D. Pistilli. Course signals at purdue: Using learning analytics to increase student success. In *Proceedings of the 2Nd International Conference on Learning Analytics and Knowledge*, LAK '12, pages 267–270, New York, NY, USA, 2012. ACM.
- [4] R. S. Baker, D. Lindrum, M. J. Lindrum, and D. Perkowski. Analyzing early at-risk factors in higher education elearning courses. In *Proceedings of the 8th International Conference on Educational Data Mining*. Educational Data Mining, November 2015.
- [5] N. Cliff. Dominance statistics: Ordinal analyses to answer ordinal questions. 114, 1993.
- [6] A. Hershkovitz and R. Nachmias. Learning about online learning processes and students' motivation through web usage mining. 5, November 2009.
- [7] S. M. Jayaprakash, E. W. Moody, E. J. Lauria, J. R. Regan, and J. D. Baron. Early alert of academically at-risk students: An open source analytics initiative. *Journal of Learning Analytics*, 1(1):6–47, 2014.
- [8] Y. Levy and M. M. Ramim. A study of online exams procrastination using data analytics techniques. 8, 2012.
- [9] M. D. Milliron, L. Malcolm, and D. Kil. Insight and action analytics: Three case studies to consider. *Research and Practice in Assessment*, 9, 2014.
- [10] StudyMode. Eighty-seven percent of high school and college students are self-proclaimed procrastinators. <https://goo.gl/C0XdZQ>, May 2014.