Scientia
Publica
Media

🔓 **Research Article**

Student Engagement as a Factor of Successful Online Course Completion

Pascal Roubides[1] ✉ iD **, Adwyn Roubides**[2] iD

[1]Department of Mathematics, Broward College, Ft. Lauderdale, FL, USA
[2]Brentwood Academy, Brentwood, TN, USA

Abstract
The following monogram discusses a quantitative case study funded via a grant award which was aimed at discovering correlations of course access levels and academic success in an online gateway mathematics course at a higher education institution in south Florida, USA. In particular, the study aimed at investigating whether student engagement in a general education online mathematics course, College Algebra, as measured by the frequency of student course logins, might correlate with successful completion of the course itself. The study addresses the need for quantitative measures to identify at-risk student behaviors in online learning environments, where subjective faculty judgment often delays or limits effective interventions. Results obtained show a positive correlation between the frequency of student course logins and their respective final course outcomes. Despite challenges accessing a higher level of detailed course engagement data, as initially intended, the study results statistically support the hypothesis that increased course access correlates with better academic outcomes. The study also aligns with activity theory, which emphasizes that meaningful learning arises from active engagement, as well as with recent studies on student engagement challenges in online learning. It is the author's point of view that the current study is an important first step in statistically quantifying such expected relations and promoting further study and application of quantitative metrics in regards to addressing online students' course engagement challenges and hopefully lead to early warning systems that can increase timely interventions, and by extension, increase overall rates of student retention and success in online environments.

Keywords: Factors for Success Online, Online Course Completion, Online Learning, Student Engagement

1. INTRODUCTION

At the present and at most higher education institutions, faculty must carry out the identification of individual at-risk behaviors in their online classes manually based on their own subjective perception of what level of online student engagement may be considered at-risk behavior rather than having access to quantitative measures in support of such perceptions. Despite the fact that online classes assume that technology is being used at all levels of their delivery, there are still many functions for online faculty that are subjective and, in the context of intervention when students may be at risk of dropping out or failing their courses, the intervention, if one does occur, may be ineffective, untimely, or even useless.

Traditionally, because the perception or timing when faculty may observe some adverse behavior, especially in accelerated classes (such as 4-, 6-, or 8-week classes), and the time it takes to contact support services, which in itself may take days and/or weeks prior to the attempt to reach and counsel a student, makes the entire collaborative effort ineffective. Thus, there is a clear need for better indicators of at-risk behavior that can accelerate the process of intervening and providing corrective actions, suggestions, or strategies to at-risk online students.

The study described herein is an analytics study with the objective to use statistical analyses in order to discover correlations between levels of engagement in online gateway mathematics courses and successful course completion (gateway mathematics courses are courses considered college credit-earning, entry-level, and/or required for further program completion, and usually suffering from low success rates). To the author's knowledge, no such study has ever taken place at the host institution (serving over 50% of its 60000 students annually in the online environment) that directly aims to predict at-risk engagement levels of online students.

It is believed that course engagement positively correlates with course performance. This is an expectation which stems from much empirical observation and the author's experience with online education. However, there are seemingly not many studies confirming this expectation. Therefore, the current study serves as a "proof of concept" to show that, statistically, student engagement with an online course positively correlates with course performance and if so, identify levels of engagement that would signify an at-risk student behavior, thus acting as an early-warning quantitative measure for initiating appropriate institutional interventions.

The study described herein attempted to discover statistical correlations of metrics, such as course engagement and the levels of that engagement, and academic success in an online gateway mathematics course. Being able to identify levels of engagement that would signify an at-risk student behavior can serve as indication of an early-warning quantitative measure for initiating appropriate and timely interventions to accomplish two important functions:

i)   Minimize the level of student attrition in the online gateway course under study.
ii) Maximize the number of students that successfully complete their online course under study.

This research can be expanded to any course of any discipline and a large-scale institutional-wide study could result in an in-house predictive model to identify engagement levels quantifying at-risk behavior for online students, thereby allowing for early intervention measures to take place in a timely manner.

1.1. Literature Review

Even though online education has been increasing in popularity, especially during the past decade, attrition remains a problem faced by many educational institutions and has been so for a long time (Bowden, 2008). This issue has persisted since the early works of Tinto in the 1970s without seemingly any improvement made to-date. Quite often, final course success is unrelated to knowledge and skill but instead it is a reflection of a lack of persistence by the student or lack of and/or weakness in related self-regulatory traits, such as motivation, organization, or time management. There has been a plethora of publications about the "best" teaching methods for the online education environment (Hart, 2012), however, little is known about how to identify students who are at risk of dropping out of an online course or being disengaged and therefore become unsuccessful (Kerr, Rynearson, & Kerr, 2006; Liu, Gomez, & Yen, 2009).

As Hart (2012) discusses, in general, students report being satisfied with the online environment and it is accepted that learning outcomes in online courses are also similar to those of the traditional classroom; however, challenges still exist which can result in an inability to complete a gateway course and, in turn, an inability to complete the student's intended program of study. This fact is well supported in literature (Ivankova & Stick, 2007; Levy, 2007; Müller, 2008; Park & Choi, 2009). Several studies have been conducted to determine factors that may be positively related to student success (Harrell & Bower, 2011; Levy, 2007). Moreover, several other studies have also assessed which factors interfere with success and how students' attitudes are related to course completion (Holder, 2007; Müller, 2008; Park & Choi, 2009).

In general, despite motivating factors for success, personal resolve and determination day after day would contribute significantly to course persistence; students who exhibit increased self-regulatory capacity, self-efficacy and have high personal expectations, as well as those who enjoy the challenge of online learning will also tend to be more persistent and lead the student to engage more in the classroom, thereby having great potential for success (Roubides, 2018). However, what student behaviors in an online course can be considered at-risk behaviors is not easily quantified and thus little is known on that front (Hart, 2012; Roubides, 2019).

This study addresses the gap in education literature in attempting to identify at-risk behaviors a-priori that can be used to provide appropriate interventions to retain and assist online students in completing their courses successfully.

2. METHODS

The research question for this study is as follows: Does course and course component access and engagement directly correlate with successful completion of an online gateway mathematics course?

This is an important question to explore because it is directly related to improving functions that go beyond the classroom, such as being able to maximize efficiency in student support services and allow for faster and more accurate identification of at-risk levels of engagement of online students, thus, allow for faster and more on-point collaboration between faculty and support personnel to address arising issues. The impact of the study's findings may be of utmost importance to all stakeholders, faculty, students, and administrators, since it was aimed at providing quantitative measures for minimizing attrition of online students while maximizing their chances for successful course completion.

The study described herein was based on a descriptive correlational quantitative design. This was deemed the most appropriate research design to employ, given that there was no need to change the environment or manipulate any variables, and no intervention was required, as well as no cause-and-effect conclusions were sought-after. As Fischer et al. (2023) indicated, such research designs are best when the goal of the study is to identify issues with or justify current practices, inform changes required with such practices, or simply develop theories about such practices.

The study used anonymous data from one section of a College Algebra course (designated as MAC1105) offered fully online during Fall semester 2024. MAC1105 is an online gateway mathematics course taught in various formats, and the format chosen for this study was the accelerated (8-wk) OER course format. The sample size was expected to be approximately 37 students (which is the currently maximum class size), with the actual sample size being 33 (after initial dropouts and withdrawals). Data used included course access frequencies of individual students and related final course outcomes (designated as success or failure). The statistical analysis required was conducted in a commercially available statistical software package and the expected time frame for completion for the study was approximately twelve weeks (eight weeks for the duration of the course and four weeks for the ensuing data collection and analysis and final reporting of results).

Data collected was data provided by the online course learning management system (LMS), which in this case was D2L. The data collected was expected to involve course access frequencies as well as course component access frequencies for each student enrolled in the target course. Data collected was expected to fall under three main categories:

i) Course access - this data simply reflects the frequency with which each student accesses the course without any reference to whether any other content within the course was accessed.
ii) Content access - this data reflects the student engagement with the course content and includes frequency as well as duration of the engagement.
iii) Performance - this is the data that all other data will be tied to; it indicates the final grade received by each student in the course.

Due to administrative and technical difficulties with the second category of expected data collection, this category of data was excluded from the ensuing statistical analysis in order to meet the study's projected timeline and the funder's final reporting deadline requirements. The remaining categories of individual student-course engagement data collected was then analyzed against related course success data for each student. The data was anonymized first to preserve privacy of the students involved and thereafter organized and structured appropriately to conduct a correlation analysis.

The independent variable in this study was a quantitative measure representing each student's course access frequency represented by each student's total number of course logins during the course, and the dependent variable was a dichotomous variable taking the value of 1 if the student completed the course successfully as defined by the host institution (i.e., passing the course with a grade of A, B, or C) or the value of 0 if the student completed the course unsuccessfully (with a grade of D or F).

It is important to note here that a grade of D is academically considered "passing", however, most US institutions require grades of C or higher for enrollment in subsequent courses, higher level courses, or for administrative and financial aid purposes. Thus, most related research considers grades of D as "non-passing".

3. LIMITATIONS

As with all research studies, there are several factors that affect the ability to generalize results obtained. First and foremost, this study can be categorized as a case study given that it was conducted on data from a single course section which was in itself limited to a maximum number of students. As such, the sample size used can be considered a limitation. Given the rather small sample size, the students in the section used for this study, may not have been entirely representative of the population of the institution and even possibly not entirely representative of the subset of students who have to take this course. However, this was only a preliminary study and one of the goals was to understand if a larger study may be necessary.

In addition to the limitation of using data from a single course, the course also was facilitated by the researcher and author of this monogram. This may also be a limitation in the sense that students in the course may have had slightly different behaviors than normal based on the pedagogical approach used, the communication style, even grading of classwork, and other characteristics that are unique to the facilitator that can be considered "researcher bias."

Moreover, as stated in the previous section, the initial intent of the study was to use data that reflected a broader aspect of student engagement with the course content that would include both frequency of logins as well as duration of the logins. Due to administrative and technical difficulties with accessing duration of login data, this category of data was excluded from the ensuing analysis in order to meet the study's projected timeline and the funder's final reporting deadline requirements. The lack of this additional data is a limitation in the generalization of results obtained.

4. RESULTS AND DISCUSSION

Data collected on the 33 students enrolled in the study's subject course showed that 28 students or 84.8% of the class completed the course successfully (shown in related tables and figures as "success") with a grade of A, B, or C. The remaining 5 students representing 15.2% of the class (shown in related tables and figures as "failure") did not complete the course successfully and were awarded a grade of D or F. These rates and relative frequencies and percentages are shown collectively in Table 1 and in graphical mode in Figure 1.

Table 1. Success/Failure Rates of Students in the Subject Course

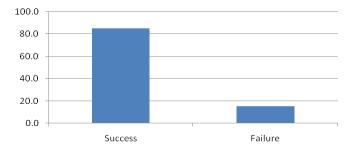|	f	f/n	%
Success	28	0.848	84.8
Failure	5	0.152	15.2

Figure 1. Percentage Rates of Success/Failure in the Subject Course

The number of logins and the students' course outcomes data was organized and grouped based on the students' final course outcome (success or failure) and used for the statistical analyses necessary to investigate the study's research question. Some descriptive statistics for the two groups of data (success and failure) in terms of the number of logins for students in each group are shown in Table 2. It is immediately apparent that the students who completed their course successfully had a much larger average number of logins in the course than those who did not complete the course successfully (but also a larger standard deviation as well). Whether such an "apparent" disparity is statistically significant is the subject of the ensuing statistical analysis.

Table 2. Mean and Standard Deviation of the Number of Logins for Each Student Group

	Mean	sd
Success	67.0	23.182
Failure	12.8	7.676

In order to establish a possible relation between the number of logins and the final course outcome in this course, Pearson's correlation coefficient r was computed. This coefficient was found to be 0.6649 which represents a moderate correlation between the two variables. The regression table for the expected value of the dependent variable (success or failure) shows a more accurate prediction of the correlation. Furthermore, the regression equation can also be used to model the relationship between the number of logins and the expected outcome using a simple linear model (this can be also seen in the Bland-Altman plot in Figure 2).

Table 3. Regression Table for the Expected Value of the Dependent Variable

	B	s.e.	t	p
Logins	0.008	0.002	4.6982	$p < 0.0001$
Intercept	0.356			

To further investigate the absence of bias, heteroscedasticity, and trend, a statistical analysis known as Bland-Altman analysis or Tukey mean-difference was used (results are shown in Table 4) and a graphical representation of its findings was generated (Figure 2). This analysis allows researchers to look for some important issues that can affect correct interpretation of results such as:

i) bias, that the difference between the two measurements is different from zero,

ii) heteroscedasticity, that the clout of points widens or narrows depending on the value on the x-axis, indicating that the reliability of a measurement changes with the value of the measurement, and

iii) trend, which indicates that the bias between the two measurements changes with the value of the measurement.

Table 4. Results of the Tukey-Bland-Altman Analysis

Tukey-Bland-Altman table				
	value	sd	95% C.I.	
Mean	58.545	28.109	48.955	68.136
Mean - 2sd	2.3269	48.687	-14.284	18.938
Mean + 2sd	114.76	48.687	98.153	131.38
bias: 58.5; t = 11.965, df = 32; $p < 0.0001$				

Based on the Tukey-Bland-Altman analysis, no significant issues were present. The related plot does confirm a linear trend between the number of logins and the final course outcome. Hence, one can safely assume that there is a positive relationship between the number of logins a student makes during the duration of the course (possibly also signifying increased engagement with the course content and activities) and the student's final course outcome.
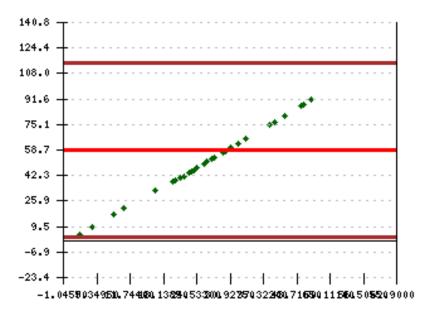
Figure 2. Tukey-Bland-Altman plot

Finally, a t-test was conducted to ensure that the apparent difference in the means of the groups of students who completed the course successfully or not was in fact not zero (in a statistical sense). In this test, the null hypothesis that the difference in the means of the groups of students who completed the course successfully or those who did not is tested against the alternative hypothesis that the difference in these group means is not zero. Those means were shown earlier in Table 2 and the results of the t-test that are shown in Table 5 do verify the correct interpretation. The p-value of the test was found to be very close to zero ($p < 0.0001$) which dictates rejecting the null hypothesis. In other words, the result obtained provides an indication that the means of the logins of the two student groups shown in Table 2 are in fact significantly different. In addition, a one-tail test (not shown here) was also conducted providing further indication that the mean of logins of the "success" group was statistically greater than the mean of logins of the "failure" group.

Table 5. Results of a T-Test for the Difference of Means Shown in Table 2

Difference	-54.2
Standard error	10.589
95% CI	(-75.8, -32.6)
t-statistic	-5.119
df	31
Significance level	$p < 0.0001$

5. FURTHER DISCUSSION

The study presented herein set out to investigate whether course and course component access and engagement directly correlate with successful completion of an online gateway mathematics course. Given that there is no knowledge of such study ever been conducted at this institution, it is considered a successful first step in correlating statistically the empirical expectation that students who access their courses more frequently (and hence possibly also engage with the course content and activities more frequently) have greater potential to successfully complete the course than those students who do not access their courses as frequently. It is important to also note though that the data analyzed in this study only accounted for the total number of logins in the course by students and not actual times of being logged in or any specific access frequencies or times of engagement with course material. Though the number of logins and the actual engagement or time spent in the course (either studying, discussing, or taking assessments) may be related, that relationship or its strength was not a subject of this study (and not feasible given data access challenges) but it does pose an interesting extension of the topic for further investigation.

Moreover, the study's finding corresponds to the basic framework of the activity theory (Kaptelinin & Nardi, 2017). A foundational aspect of the activity theory is that activity produces meaningful learning and is not an antecedent to learning. Activity theory tells us that the process of learning involves interaction with the course content, instructor, peers, and course related activities. The most common unit of analysis for activity theory is any specific course activity, which is the unit of analysis being examined by students' total number of logins in this study. Given this framework, it would be logical to expect that those students in this study who were more "active" achieved a higher potential of success as measured by students' final course outcomes.

The study can be extended in many different ways to ensure better potential for generalizing results obtained. This can be achieved by including:

i) more sections of the same course facilitated by the researcher,

ii) sections of the same course facilitated by other instructors,

iii) sections of all general education mathematics courses,

iv) sections of any mathematics courses at the college.

The potential impact this initial study and/or extending this study may have is of utmost importance in the pursuit of college-wide improved rates in metrics such as retention and success of our students. It is believed that results obtained from this study can be replicated in different sections of the same course, different courses, as well in courses of all disciplines, and results can be used for the benefit of future online students in terms of being able to identify potential at-risk behavior early enough, especially in courses delivered during accelerated terms, thus allowing for sufficient and timely interventions to be implemented. The above is also supported by a recent systematic review by Akpen et al. (2024) during which eighteen studies spanning the years 2019-2024 were examined and results were aggregated for the authors' report. Results and implications discussed in the current limited study seem to lead to similar conclusions, namely, that maintaining student engagement is always a challenging issue and therefore any progress toward identifying ways to provide information of at-risk behavior as early as possible can be crucial in addressing the issue.

6. CONCLUSION

The study presented in this monogram was a descriptive correlational quantitative case study. It emanated from the author's impetus to provide a starting point by which further studies can verify statistically the general expectation that increased frequency of logins in an online course can lead to potentially increased engagement with the course and its components, thus resulting in increased success rates of course completion. More specifically, the study aimed at investigating whether student engagement in a general education online mathematics course (MAC1105 College Algebra), as measured by the frequency of logins in the course, may correlate with successful completion of the course itself.

Despite any limitations in the study, results obtained show a positive correlation between the frequency of logins in the course under study and successful course completion. The study supports the hypothesis that increased course access correlates with better final course outcomes. The study aligns with related theories (activity theory in this case), which emphasizes that meaningful learning arises from active engagement. The study results are also supported by similar studies in current literature which express the sentiment of a factual challenge to maintain student engagement in online courses, and that being able to identify at-risk behaviors is a crucial step in addressing the issue. From the author's point of view, this case study represents an important first step in statistically quantifying such a relation and promoting further study and application of quantitative metrics in regards to online course engagement and its impact on student retention and success.

Conflicts of Interest. There are no conflicts of interest to declare.

Funding. This research received no external funding.

REFERENCES

Akpen, C. N., Asaolu, S., Atobatele, S., Okagbue, H., & Sampson, S. (2024). Impact of online learning on student's performance and engagement: a systematic review. *Discover Education, 3*(1), 1-15.

Bowden, J. (2008). Why do nursing students who consider leaving stay on their courses? *Nurse Researcher, 15*(3), 45-59.

Fischer, H. E., Boone, W. J., & Neumann, K. (2023). Quantitative research designs and approaches. In *Handbook of research on science education* (pp. 28-59). Routledge.

Harrell, I. L. & Bower, B. L. (2011). Student characteristics that predict persistence in community college online courses. *American Journal of Distance Education, 25*(3), 178- 191.

Hart, C. (2012). Factors Associated with Student Persistence in an Online Program of Study: A Review of the Literature. *Journal of Interactive Online Learning, 11*(1), 19-42.

Holder, B. (2007). An investigation of hope, academics, environment, and motivation as predictors of persistence in higher education online programs. *The Internet and Higher Education, 10*, 245-260.

Ivankova, N. V., & Stick, S. L. (2005). Collegiality and community-building as a means for sustaining student persistence in the computer-mediated asynchronous learning environment. *Online Journal of Distance Learning Administration, 8*(3).

Kaptelinin, V., & Nardi, B. (2017). Activity theory as a framework for human-technology interaction research. *Mind, Culture, and Activity, 25*(1), 3-5.

Kerr, M. S., Rynearson, K., & Kerr, M. C. (2006). Student characteristics for online learning success. *The Internet and Higher Education, 9*(2), 91-105.

Levy, Y. (2007). Comparing dropouts and persistence in e-learning courses. *Computers and Education, 48*, 185-204.

Liu, S. Y., Gomez, J., & Yen, C. (2009). Community college online course retention and final grade: Predictability of social presence. *Journal of Interactive Online Learning, 8*(2), 165-182.

Müller, T. (2008). Persistence of women in online degree-completion programs. International *Review of Research in Open and Distance Learning, 9*(2), 1-18.

Park, J. H., & Choi, H. J. (2009). Factors influencing adult learners' decision to drop out or persist in online learning. *Educational Technology & Society, 12*(4), 207-217.

Roubides, P. (2018). *Self-regulation and Successful Completion of Mathematics Online*. 25th Int'l Conference on Learning, Athens, Greece.

Roubides, P. (2019). *Improving Retention & Success of Online Learners via Front-End Behavioral Analytics*. 11th International Conference on Education and New Learning Technologies, Palma de Mallorca, Spain.