6-2018

How Engaged are our Students? Using Analytics to Identify Students-at-Risk

Anthony Williams
Avondale College of Higher Education, tony.williams@avondale.edu.au

Jason K. Morton
Avondale College of Higher Education, jason.morton@avondale.edu.au

Peter W. Kilgour
Avondale College of Higher Education, peter.kilgour@avondale.edu.au

Maria T. Northcote
Avondale College of Higher Education, maria.northcote@avondale.edu.au

Recommended Citation

Williams, T., Morton, J., Kilgour, P., & Northcote, M. (2018). How engaged are our students? Using analytics to identify students at-risk. In T. Bastiaens (Ed.), *Proceedings of EdMedia: World Conference on Educational Media and Technology* (pp. 122-128). Amsterdam, Netherlands: Association for the Advancement of Computing in Education (AACE). Retrieved from https://www.learntechlib.org/p/184189

How Engaged are Our Students? Using Analytics to Identify Students At-risk

Tony Williams
Jason Morton
Peter Kilgour
Maria Northcote
Avondale College of Higher Education
Australia
tony.williams@avondale.edu.au
jason.morton@avondale.edu.au
peter.kilgour@avondale.edu.au
maria.northcote@avondale.edu.au

Abstract: Learning Management System (LMS) analytics have become an area of increasing interest and development. The potential to better understand our students' levels of engagement provided by the systems have, to date, has been underutilized information resources. The study reported here looks at the relationship of student and staff engagement in the LMS and considers the levels of predictability in student behavior leading to failure. Also considered is the impact of the lecturer on the student engagement of poor and high performing students.

Introduction

Analytics drawn from an institution's Learning Management System (LMS) has to date been an underutilized resource as they are able to provide significant information regarding students but, in general, these data are not well understood or used to full advantage. The analytics can provide the opportunity to better support our understanding of student behaviours, especially the extent to which they engage with their learning. It is proposed that student success in a subject is likely to positively correlate with their level of engagement. This paper reports on the early phases of a project at Avondale College of Higher Education in Australia, a small college of approximately 1200 students. Avondale provides a good opportunity to better understand student and teaching staff engagement using the analytics drawn from Moodle, Avondale's LMS. This investigation provides an opportunity to identify ways to support students, especially those most at risk of failing, which will ideally lead to higher student success rates and/or higher completion rates.

The motivation for an institution to have high success or completion rates has two primary dimensions: 1) to provide a more supportive learning environment for students, and 2) financial reasons. Institutions therefore have an interest to ensure students succeed despite a climate where many students fail to do so. For example, Pitkethly and Prosser (2001) proposed that one third of all university students contemplate withdrawing during their first year of study. Some of the reasons for withdrawal are likely linked to the tensions experienced by university students as identified by Krause (2005). These include:

- the relevancy to themselves of the program they are enrolled in;
- perceptions of themselves as clients (from the marketing and service dimensions of their institution); and
- the disciplinary and academic integrity standards required by academics.

There have been numerous approaches aimed at reducing attrition including increasing levels of student engagement, creating learning communities, and improving tactics to construct academic and social integration. These have been shown to have a positive impact on student retention (Tinto & Goodsell-Love, 1993; Zhao & Kuh, 2004). In more recent times the potential of analytics drawn from an institution's LMS is seen as a way of identifying "at risk" students who are "on track" to fail. The ability to support students from an informed position of their engagement will not only assist students but also potentially reduce the number of students who withdraw from university studies, thus reducing significant cost to the student as well as preventing future income losses to the university. The LMS, a technology used by staff and students and which has the ability to track their usage, has the potential to provide insight into patterns of engagement allowing for a more strategic approach to student support and encouragement.

This paper explores the predominantly untapped, but increasingly viable potential of LMS analytics to inform ways of better supporting and encouraging students, including identifying critical times when this support and encouragement is most needed.

Literature Review

Universities endeavor to improve their student retention by researching student experience with a view to improvement. Supporting a student in their learning has many benefits in improving student retention, with the bonus of increasing graduate quality. Through the use of analytics there can be a focus on identifying students with low engagement profiles, then these students can be monitored and encouraged, and provided with direction or support to assist them in greater engagement with their learning.

Considerable evidence suggests a direct correlation between students' level of engagement and their final grades. For example, Kim, Park, Yoon, and Jo (2016) identified a correlation between levels of student activity and engagement in online discussions and their final subject grades, specifically with reference to the domains of active participation (total time in LMS/discussion, frequency of LMS/discussion visit, number of postings), engagement with discussion topics (posting length, discussion time per visit), consistent effort and awareness (regularity of visits and time lapse between visits), and interaction (number of responses triggered by a post, number of replies to received responses).

This has led to an interest in the early identification of students likely to be at risk of poor performance. In the case of Zacharis (2015), a practical model was developed for predicting students that are unlikely to succeed in blended learning subjects. Zacharis analyzed usage data stored in the log files of LMSs which allowed the development of timely, evidence-based interventions to support at-risk or struggling students. An earlier study of early intervention strategies looked to identify and support students at-risk, students who have failed too many subjects to be allowed to continue in their studies (Williams & Sher, 2007). The Williams and Sher study utilized analytics as part of a range of protocols had the potential to identify students who were having difficulties in their studies and through provision of strategic support for students. While these strategies need further development, current work is drawing on the significant capacity of LMS analytic systems which, until now, have been largely underutilized. These systems that record student use of LMSs through tracking and analysis of online data (analytics), have the potential to support staff in identifying poor performing students.

You (2015) investigated the capacity of LMS data to identify self-regulated learning and assessed the association between self-regulated learning data and course success. In recognition that LMSs capture large amounts of information about the frequency, patterns and sessions of digital learning activity, You identified which indicators significantly predicted course achievement and whether mid-course data could sufficiently predict final subject outcome. The study involved 530 university students across 13 online subjects and used data drawn from factors such as study regularity, total viewing time, number of sessions, number of late submissions, proof of reading course material, and messages created. You concluded that teachers need to encourage students to engage regularly with the course materials, to use social media as well as the LMS to communicate with each other, and that procrastination pre-determines course achievements.

Looking further, Agudo-Peregrina, Iglesias-Pradas, Conde-González, and Hernández-García (2014) outlined three "system-independent" interaction classifications: those based on the agent (student-student, student-teacher, student-content); those based on frequency of use (Most Used—transmission of content; Moderately Used—discussions; student assessment/evaluation; and Rarely Used—subjects/teacher/satisfaction evaluation surveys, computer-based instruction); and classification based on participation mode (active vs. passive interaction). They evaluated the relationship between each component and academic performance across two different learning modalities: blended learning and online learning, and developed an extraction and reporting plug-in tool for the LMS to automatically classify interactions into the appropriate category. Results indicated that student performance correlated with active interaction across each. Recent advances have also been highlighted by Cerezo, Sánchez-

Santillán, Paule-Ruiz, and Núñez (2016) in the use of educational 'data mining' in LMS and how it can assist, identify, and predict students' learning styles, effort expended, and learning achievement.

In another study, by Strang (2016) correlation analysis was used to investigate whether a predictive relationship exists between student grades and student age, gender, culture, and LMS engagement. The only significant predictor of student grade was LMS log-in frequency and engagement in online assignments. Though a relationship between LMS engagement and student success at the assessment item level was found, this was limited to using log-in frequency as the data source, which is not a complete measure of student engagement in an LMS.

From the insights of the studies cited above there is strong evidence that student engagement levels with a subject's LMS does correlate with success, though a question remains as to whether the levels of student engagement early in a semester can be used to predict poor student performance. Also, of interest is the relationship between engagement with the LMS by teaching staff and student engagement. The study reported in this paper examines these issues to gain a better understanding of the link between students' engagement with their LMS and their performance, and if the engagement of teaching staff has an impact on the engagement of students.

Methods

Analytic data were collected from Avondale College's LMS for eight subjects conducted in 2017 in the disciplines of Theology (n=4), Education (n = 2), Arts (n=1) and English (n=1). All subjects were managed by staff with an active LMS presence and were chosen to represent a diversity of teaching styles and year levels. All subjects ran concurrently in the same semester. All LMS activity of staff and students, recorded as the time and date of user selections or 'clicks', was acquired then managed within Excel. This large data set was filtered to remove the activity of non-teaching staff and those students who withdrew from each subject, and also any activity that occurred outside the 13-week teaching plus one-week exam period. This reduced the data set to 127 student users and nearly 63,000 activity logs.

The data were then sorted by student name, and the overall subject grade (i.e., Fail, Pass, Credit, Distinction or High Distinction) was entered for the activity logs of each student. The data were then re-sorted by date and each date was allocated to a two-week teaching period (i.e., 1 & 2, 3 & 4, 5 & 6 etc.). Pivot tables were used next to determine the total activity logs recorded in each two-week period for each grade, and these were then converted to averages per student before being graphed. Average activity logs in each two-week period were also determined for teaching staff.

Interpretation of the outcomes was taken from the trends observed on the line graphs. T-tests were conducted to evaluate the significance levels of the differences between the number of LMS hits for students with different grade outcomes. Taken in context, the visual trends from the graphs of the times in the semester when students were accessing the LMS were of prime concern in this study. A Pearson Product-Moment Correlation was used to substantiate these different patterns of LMS access.

Results

While there appeared to be no conclusive pattern of LMS activity logs for students who achieved a Pass, Credit or Distinction grade (Figure 1), there was a distinct and revealing pattern of usage when a contrast was made between those with a Fail grade and those who received a High Distinction (Figure 2). When comparing the activity logs for students receiving Fail and High Distinction grades, the following can be observed:

- The number of activity logs of a High Distinction student is significantly higher ($p=0.0003$) than for students who received a Fail grade.
- Students who achieved a High Distinction grade demonstrated consistent LMS usage between Weeks 1 and 6 setting them up for a more informed second half of the semester. Indeed, High Distinction students demonstrated a dramatic upturn in activity logs in the same period.
- Failed students demonstrated a pattern of low usage at the start of the semester that dropped away even further from Weeks 5 to 8 which is the period of the semester when many assessments are due.

- As final assessments and examination time approached around Week 10, High Distinction students reached a climax in their usage of the LMS that set them up for the final approach to the end of the examination period. Students with a Fail grade appeared to leave their major access to the LMS until the end of their semester. In fact, their access peaks in the examination period.

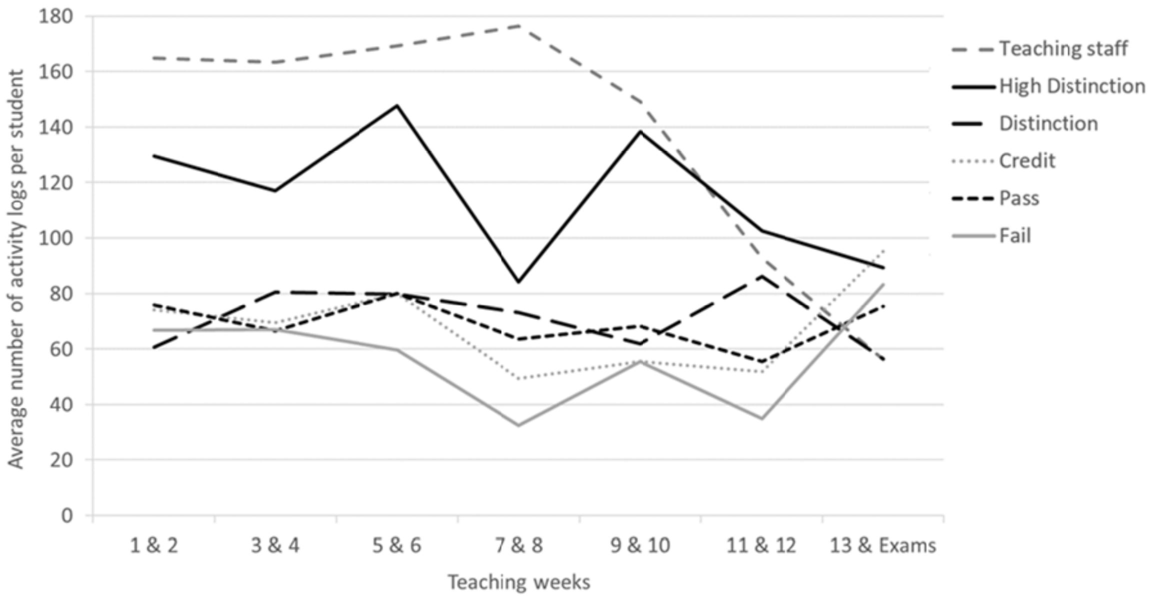

Figure 1: Activity logs on LMS for teaching staff and students showing all achievable grades.

The relationship between LMS activity logs of teaching staff and students is also highlighted in each figure. It can be seen that at the most intense times in the semester, the activity of teaching staff and High Distinction students is very similar while the usage for students with a Fail grade is far below. These intense times are Weeks 5 and 6 when many high stakes assessments are typically due and again in Weeks 11 and 12 when final assessments and examinations are being prepared for and the lecturer is posting material to help with these tasks.

A Pearson Product Moment Correlation was calculated to show the relationship between the trends in the LMS access of both groups of students with that of their lecturer. It was found that the correlation between the activity logs of students with a Fail grade and teaching staff was $r = -0.30$ indicating a negative relationship, whereas for students who obtained a High Distinction grade the coefficient was $r = 0.47$ which indicates a moderate positive relationship. This illustrates that, unlike the High Distinction students, the students with a Fail grade are not following the activity of the teaching staff, whether that be the provision of resources or other support.

Discussion

The findings of this study support a preliminary study (Williams, Kilgour, Stewart, & Northcote, in-press) conducted in semester one of 2017 which was limited to a single subject that was a combination of traditional face-to-face and blended learning, and was undertaken using data obtained at only three times during the semester: Weeks 5, 10 and 14. The preliminary study was narrow in scope and was primarily a *proof of principle* for the study reported here. The preliminary study did identify a close relationship between student engagement, measured in click counts at Week 5, and the students' final grades. The study reported in this paper extended the scope of the earlier study, and incorporated a larger cohort of students in multiple subjects, a range of levels and different teaching modes. The study also looked at student engagement levels at fortnightly intervals to establish the relationship between students' engagement and their ultimate grade. The study also looked at the impact that the online presence of teaching staff had on the online behavior of students.

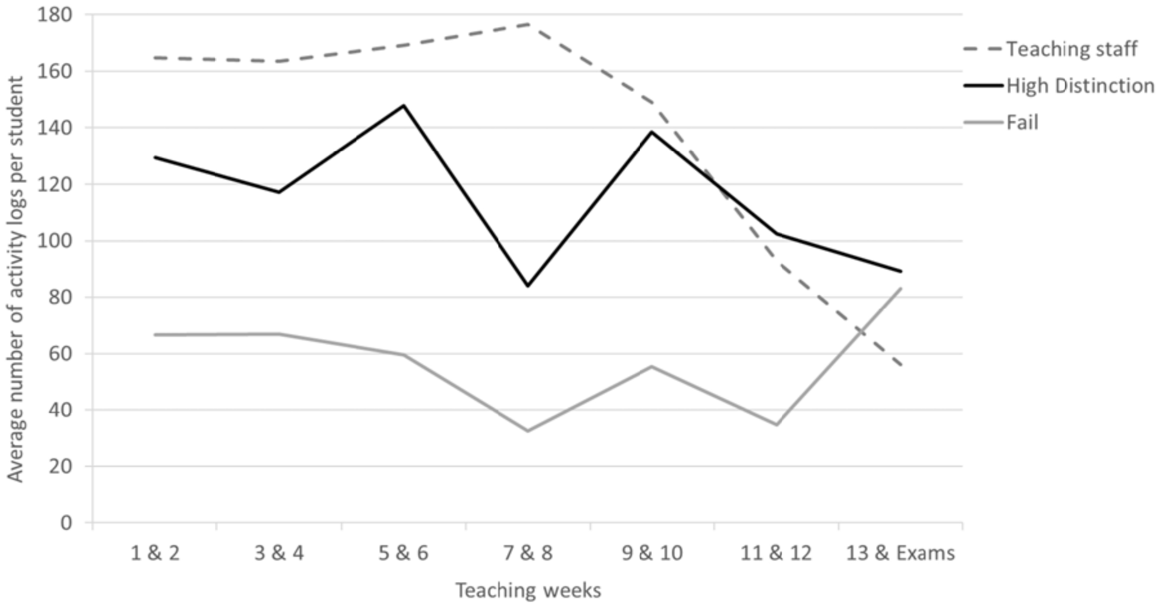

Figure 2: Activity logs on LMS for teaching staff and students showing High Distinction and Fail grades.

Although this study involved limited student numbers, it did provide some interesting insight into the students' online behavior in the subject's LMS and the students' success rate in the subject. A strong correlation was found between student engagement in the first six weeks of a semester and the ultimate outcome for the students. Students who had not engaged by Week 6 were a distinct proposition for failure in the subject, no matter how much they increased their engagement after Week 6. This highlights the importance of strategies designed to promote student engagement in these first six weeks.

The findings of this study are consistent with the previous work of You (2015), Agudo-Peregrina et al. (2014) and Strang (2016), but the present study brings the dimension of time into consideration. Week 6 appears to be pivotal in terms of the work done by students on the LMS up to that time. It is possible to start to put into practice some initiatives which would utilize the forecast for the purpose of intervening where a student is identified as a potential Failure and contact the students who are exhibiting poor engagement and encourage them to engage more fully, though it would be good to identify the trend earlier in the semester. More work needs to be done in this area to improve predictability earlier in the semester which may involve the consideration of other measures such as rates of subject outline download and dwell times. Student engagement with links to the library's resources in the LMS also needs to be monitored.

The engagement of teaching staff in the subject's LMS was interesting as it did not impact to a great level, on the engagement levels of students who achieved a Fail grade. The fact that the students who achieved a High Distinction respond positively to the engagement of teaching staff is an interesting finding, but as yet not fully understood. The type and prevalence of teaching staff engagement that impacts student engagement needs to be established, along with investigation into how their engagement can be modified to better support less-engaged students. More work is needed in this area with more defined study protocols to look at student responsiveness to the engagement of teaching staff.

One interesting insight that came from this study was the behavior of some students who went on to fail the subject. Some students who failed had the highest rate of engagement with the LMS, totally divergent from the typical profile of the Fail student. This behavior was difficult to understand. To further understand this phenomenon, the lecturer was interviewed in an attempt to establish reasons for this behavior. It was identified that there were two possible scenarios for this:

- The subject was not a fully online/blended subject and class attendance was necessary. The students with high engagement levels were those that did not attend class but tried to pass by gaining everything from the LMS.
- The students had very poor technique and were just logging on and clicking through the materials without genuinely engaging.

Learning analytics can breach the uncertainty around how to allocate resources, develop competitive advantages, and most importantly, improve the quality and value of the student learning experience. So, there is much work to be done in better understanding student engagement and how it can be measured and used to assist teaching staff. Initiatives for the use of learning analytics include:

- For educators, the availability of real-time insight into the performance of learners—including students who are at-risk—can be a significant help in the planning of teaching activities.
- For students, receiving information about their performance in relation to their peers or about their progress in relation to their personal goals can be motivating and encouraging.

Before these objectives are to be implemented, the scope of our research and development activities needs to broaden to carry out comprehensive analytics and to broaden the scope of where the data comes from. Currently the data is extracted from the LMS, but there is scope to broaden this to other systems in the institution. Various institutional systems, in particular the LMS, the student information system, and a variety of library systems could be utilized. Overall, the development of analytics can be enhanced for use by educational organisations.

Conclusion

The research findings described in this paper have contributed to the wide-ranging discussion currently underway in the higher education sector regarding the value of gathering data about students' access of online courses and LMSs, also known as *learning analytics*. While the last few years have seen learning analytics collected across multiple learning platforms, disciplines, institutions and countries, less research has been published that focuses on how such analytics have been *used* to investigate the impact of online student learning activity and non-activity on student achievement levels and learning outcomes. The results of the study reported in this paper illustrate an institution-wide example in which students' access to online learning materials, specifically via the institution's LMS, was linked to their final grade, which represented their achievement of course learning outcomes. This research has provided evidence of how the online access patterns of students from one higher education institution appear to be linked to the high and low extremes of student achievement levels: specifically, the lowest levels of student learning outcomes (reflected in Fail grades) and the highest levels of student achievement (reflected in High Distinction grades). While the results did not reveal any definitive connection between students' access patterns and their achievement of the low to medium level grades (that is, Pass, Credit and Distinction, equating to achievement scores of 50-84%), the results show a relationship between the students' access patterns to online materials and their achievement of a Fail or High Distinction grade.

Both groups, the students who achieved a Fail or High Distinction grade, demonstrated a fairly consistent level of access to the course over the semester, with a few peaks and troughs at particular stages of the semester. Students who failed the course typically demonstrated consistently low access to their online course, with especially low levels of access mid-semester. Interestingly, their highest access level occurred at the very end of the semester. Students who achieved High Distinction grades also regularly accessed the online materials, but their access levels were much higher than the students who achieved a Fail grade. The High Distinction students' access also dipped mid-semester. Interestingly, the online access patterns of students who achieved High Distinction grades somewhat mirrored their lecturer's online access pattern, especially at the beginning and end of the semester, allowing for the lag between the lecturer engaging and then the students responding. Although These findings support the allocation of resources, ideally at both the institutional and faculty level, to: a) monitor online student access to learning materials; and b) engage with students about their access, especially in the first half of the semester.

As well as adding to our knowledge of the *use* and *application of* learning analytics, rather than simply the *collection* of data generated from learning analytics, this research suggests a relationship may exist between the access patterns of students who achieve high grades and the teaching staff. More research with greater numbers of students across multiple institutions and disciplines is required to establish whether this correlation is evident in other higher education contexts.

References

Agudo-Peregrina, Á. F., Iglesias-Pradas, S., Conde-González, M. Á., & Hernández-García, Á. (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.

Cerezo, R., Sánchez-Santillán, M., Paule-Ruiz, M. P., & Núñez, J. C. (2016). Students' LMS interaction patterns and their relationship with achievement: A case study in higher education. *Computers & Education, 96*, 42-54.

Kim, D., Park, Y., Yoon, M., & Jo, I.-H. (2016). Toward evidence-based learning analytics: Using proxy variables to improve asynchronous online discussion environments. *The Internet and Higher Education, 30*, 30-43.

Krause, K.-L. (2005). *The changing face of the first year: Challenges for policy and practice in research-led universities.* Paper presented at the First Year Experience Workshop.

Pitkethly, A., & Prosser, M. (2001). The first year experience project: A model for university-wide change. *Higher Education Research & Development, 20*(2), 185-198.

Strang, K. D. (2016). Do the critical success factors from learning analytics predict student outcomes? *Journal of Educational Technology Systems, 44*(3), 273-299.

Tinto, V., & Goodsell-Love, A. (1993). Building community. *Liberal Education, 79*(4), 16-22.

Williams, A., Kilgour, P., Stewart, B., & Northcote, M. (in-press). An initiative to support first year students and students at risk: The Virtual Mentor Programme. *The Journal of Adventist Education*.

Williams, A., & Sher, W. (2007). *Using Blackboard to monitor and support first year engineering students.* Paper presented at the 18th annual Australasian Association for Engineering Education Conference, Melbourne, Australia. Retrieved from http://www. cs. mu. oz. au/aaee2007/papers/paper_34. pd f, Melbourne.

You, J. W. (2015). Examining the effect of academic procrastination on achievement using LMS data in e-learning. *Journal of Educational Technology & Society, 18*(3), 64.

Zacharis, N. Z. (2015). A multivariate approach to predicting student outcomes in web-enabled blended learning courses. *The Internet and Higher Education, 27*, 44-53.

Zhao, C.-M., & Kuh, G. D. (2004). Adding value: Learning communities and student engagement. *Research in higher education, 45*(2), 115-138.