View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by St Andrews Research Repository

Student Success on Face-to-Face Instruction and MOOCs

What can Learning Analytics uncover?

Adriana Wilde[*]
Electronics and Computer Science
University of Southampton,
SO17 1BJ, United Kingdom
agw106@ecs.soton.ac.uk

Ed Zaluska
Electronics and Computer Science
University of Southampton,
SO17 1BJ, United Kingdom
ejz@ecs.soton.ac.uk

David Millard
Electronics and Computer Science
University of Southampton,
SO17 1BJ, United Kingdom
dem@ecs.soton.ac.uk

ABSTRACT

There are fundamental differences between the face-to-face instruction model and that of Massive Open Online Courses. This paper hypothesises that despite these fundamental differences, the success factors in these learning contexts are comparable. The factors contributing to student success might be related to the same basic principles, even if manifesting themselves differently in each context - especially as the very definition of success is closely dependent on what can be measured. Learning analytics can help to uncover the indicators which have the potential for identifying students at-risk and for institutions to exercise a timely intervention.

Categories and Subject Descriptors

J.4 [**Computer Applications**]: Social and Behavioural Sciences.; H.1.2 [**User/Machine Systems**]: Human information processing.; K.3.m [**Computers and Education**]: Miscellaneous.

Keywords

Learning analytics, higher education, MOOCs, face-to-face instruction.

1. CONTEXT

Educational stakeholders seek to understand how to help students to be successful in their learning, both at a personal and an institutional level. The ultimate tool for Higher Education institutions (HEI) would be one with which at-risk students are promptly identified and given adequate, timely support, so that they can remain on track and "succeed" in their academic studies. Academic success can be approached, however, from various perspectives: student satisfaction, retention, completion, achievement and progression. Firstly, *student satisfaction*, though arguably subjective and by nature difficult to link to traditional performance indicators, can have follow-on effects on recruitment to courses, as well as retention and completion. HEI regularly conduct surveys of student satisfaction and the results are increasingly taken into account by prospective students when selecting a place to study. *Successful progression* can be argued as relevant at a personal level (e.g. a student may wish to progress to a postgraduate course on completion of their first degree), but is also relevant at an institutional level when within the institution or if the progression is regarded as that from one level to the next within a given programme. Additionally, and increasingly so, many HEI offer Massive Open Online Courses (MOOCs) to raise their profile and broaden their curriculum, which can ultimately attract highly-motivated students to progress from the taster-like, short, online course to a mainstream full course.

Retention, *completion* and *achievement* are, like the previously mentioned two aspects of success, regularly monitored by HEIs, and have a long tradition in higher education research. *Retention* is the scope to which learners persist within the HEI [3] or continue their studies at the institution (as opposed to "dropping out", which is a specially challenging problem in online learning in general and in MOOCs in particular [4]); *completion* is the rate of the annual intake of student who finish their studies on obtention of their qualification; and *achievement* is the performance as evaluated for the purposes of degree classification or similar.

There is substantial research on these approaches to measuring academic success. In particular, *learning analytics* are concerned with the analysis of data to extract characteristics of students and learning activity that could characterise student performance and offer a predictive model for achievement, typically to inform stakeholders at educational institutions [1, 5, 7, 8]. This area of research is relatively new but it has strong foundations of decades of research (mostly US-centric) [6], from where factors influencing student success as defined earlier in traditional contexts have been identified – mainly conditioning factors such as academic ability as demonstrated via admission tests and socioeconomic status, as well as whether there has been delayed entrance to the educational system (regardless of the reason). Other factors explored in the literature are the size of the institution, the existence of student loans, and the field of study chosen, all of which to greater or lesser extent, have an effect in either retention, completion or both.

[*]ACM Member 5912888.

The educational landscape is in constant change, particularly with the proliferation of MOOCs and VLEs to support F2F instruction. So, it is of interest to explore the data trail students generate navigating this landscape to gain information to a) predict success and b) offer a model for achievement. It is important to look at what elements of this form of tuition are associated with the most successful outcomes, to inform stakeholders and support decision-making for timely student interventions. Because of this, leading institutions across the world have been using learning analytics and making plans to use them to improve the quality and productivity of their operation [9, 10].

2. RESEARCH QUESTIONS

The overarching questions of concern are whether there are measurable factors for learning success that are common to the context of F2F instruction and MOOCs, and whether a parallel can be drawn despite the differences in educational context in these models of interaction. What can be learned via learning analytics? To pursue these questions, we need to look closely at data from both worlds, asking context-dependent questions to each dataset.

In particular, questions that can be asked to MOOC data are: What are the predictors of participants *completion*? Does the composition of various types of MOOC activities have a measurable effect on participants completion? What type of activity students complete/engage on the best? Does the perceived difficulty of *activities* have a measurable effect on completion?

In the F2F context these questions become: What are the predictors of student *retention*? Are there factors in the initial demographics which can predict non-completion? What type of modules students complete/engage on the best? Do past failures predict areas of future failure? Do students fail in certain modules rather than others? Is a failure on an early module more predictive of future failures than failures on later modules? Does the perceived difficulty of *modules* have a measurable effect on completion?

3. METHODOLOGY

Algorithms of machine learning implemented in the Weka toolset [2] will be applied to two datasets comprising student data. The first of these datasets is from an HEI providing F2F instruction and the other from a MOOC provider. The former contains data from three cohorts of 700+ students each enrolled at the Engineering undergraduate programme of the University of Chile (between 2010 and 2012). This course is a professional degree of 6 years of duration which includes a bachelor in engineering and sciences organised in semesters. The dataset contains student data for modules of the foundation curricular stage (the first three semesters of the degree). At this institution, programs are designed to be finished in either 10 or 12 semesters, but reportedly students typically complete their engineering programs in no less than 16 semesters, making it critically interesting for this HEI to understand the factors that influence student success. Specifically in the dataset, besides the general conditioning factors (what is known about the student during admission) related to students, there is detailed performance information in each of the modules taken (at the end of the semester but also with up to three interim evaluations), which could be used to investigate some of the questions presented in Section 2.

The remaining questions in Section 2 can be explored using the second dataset of interest, which refers to data collected the "Understanding Language" FutureLearn MOOC, provided by the University of Southampton in 2014 to 12,457 participants. This free course ran over four weeks, with an expectation of three-hours a week commitment for participants to engage in a selection of activities (videos, audios, texts, discussions or exercises). This dataset contains detailed performance information of the participant in these activities, specifically when did each activity first was attempted and last completed (if at all). A model for affective states, such as perceived difficulty of the task, frustration and boredom will be applied (similar to that in [8]).

Finally, analysing the findings from both studies will address the overarching research questions, establishing the commonalities and differences in factors for academic success in both educational contexts.

4. REFERENCES

[1] R. Ferguson. Learning analytics: drivers, developments and challenges. *International Journal of Technology Enhanced Learning*, 4(5):304–317, 2012.

[2] M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann, and I. H. Witten. The weka data mining software: an update. *ACM SIGKDD explorations newsletter*, 11(1):10–18, 2009.

[3] R. Jones. Student retention and success: a synthesis of research. *EvidenceNet, Higher Education Academy*, April 2008.

[4] Y. Lee and J. Choi. A review of online course dropout research: implications for practice and future research. *Educational Technology Research and Development*, 59(5):593–618, 2011.

[5] L. P. Macfadyen and S. Dawson. Mining LMS data to develop an "early warning system" for educators: A proof of concept. *Computers & Education*, 54(2):588–599, 2010.

[6] E. T. Pascarella and P. T. Terenzini. How college affects students: A third decade of research. *Journal of College Student Development*, 45(5):589–592, 2005.

[7] V.-A. Romero-Zaldivar, A. Pardo, D. Burgos, and C. D. Kloos. Monitoring student progress using virtual appliances: A case study. *Computers & Education*, 58(4):1058–1067, 2012.

[8] J. A. Ruipérez-Valiente, P. J. Muñoz-Merino, D. Leony, and C. D. Kloos. ALAS-KA: A learning analytics extension for better understanding the learning process in the Khan Academy platform. *Computers in Human Behavior*, 2014.

[9] G. Siemens, S. Dawson, and G. Lynch. Improving the quality and productivity of the higher education sector: Policy and strategy for systems-level deployment of learning analytics. *Australian Government: Office for Learning and Teaching*, 2013.

[10] B. Tynan and S. Buckingham Shum. Designing systemic learning analytics at the Open University. *Strategy and Policy for Systemic Learning Analytics, SoLAR Open Courses*, 2013.